**Charles University**

**Faculty of Science**

Study programme:

Anthropology and Human Genetics

**Mgr. Anežka Kotěrová**

# AGE-AT-DEATH ESTIMATION OF THE ADULT SKELETON: NEW APPROACHES TO THE EVALUATION OF SENESCENCE INDICATORS

## Odhad věku dožití dospělého jedince podle kostry: Nové přístupy hodnocení indikátorů senescence

Dissertation thesis

Supervisor:

prof. RNDr. Jaroslav Brůžek, CSc., Ph.D.

Consultant:

doc. RNDr. Jana Velemínská, Ph.D.

Prague, 2019

**Prohlášení:**

Prohlašuji, že jsem závěrečnou práci zpracovala samostatně a že jsem uvedla všechny použité informační zdroje a literaturu. Tato práce ani její podstatná část nebyla předložena k získání jiného nebo stejného akademického titulu.

V Praze, dne 11. 12. 2019

……..

Mgr. Anežka Kotěrová

# ACKNOWLEDGEMENT

Most of all, I thank my supervisor Professor Jaroslav Brůžek, to whom I am grateful for his guidance and the opportunity to travel as part of my studies. I am also very thankful to my consultant, Associate Professor Jana Velemínská, whose professional help and advice have always been very helpful.

Many thanks to all the curators of the osteological collections, without whose kind help this research would have been impossible. I thank them not only for making the collections available to me, but also for their care and hospitality during my fellowships. Namely, they were Professor Eugénia Cunha of Coimbra University (curator of the 21st Century Identified Skeletal Collection), Professor Sofia Wasterlain of Coimbra University (curator of the Coimbra Identified Skeletal Collection), and Doctor Jocelyn Desideri of the University of Geneva (curator of the Simon Identified Skeletal Collection). Last but not least, I would like to thank Professor Nawaporn Techataweewan of the University of Khon Kaen (curator of the Khon Kaen University Collection), who was in a difficult situation as curator during my stay, yet she made sure that my stay was not significantly affected by it. My thanks also go to members of the laboratories that I visited. Special thanks belong to Doctor François Marchal for lending the scanner, without whose kindness I would not have been able to digitize the material at all.

I would like to thank my colleagues from the Laboratory of 3D Imaging and Analytical Methods, as well as the other members of the Department of Anthropology and Human Genetics, who are a great team and have created a fine working environment which I always liked to work in. I am especially grateful to my friend Rebeka, with whom I disputed all the way to this point.

My thanks also belong to MSc. Lenka Červenková, who helped me especially in matters concerning my funding sources. Many thanks also go to Doctor Michal Štěpanovský for his help in the data mining field and to MSc. Nikola Pajerová who helped me with data analyses. I would also like to thank all the other colleagues with whom we share common publications.

My research would never have been possible without the financial support provided by the following funding sources:

- Erasmus+ Programme (fall 2016)
- Support for the Internationalisation of Charles University (winter 2018)

– The Hlávka Foundation – Nadace Nadání Josefa, Marie a Zdeňky Hlávkových (winter 2018)
– Grant Agency of Charles University, project GA UK No. 642218 (2018–2020)

I wish to thank whole family, especially my parents; they have always trusted me and left everything related to my studies in my hands, which I have always appreciated very much. Final thanks go to Jindra, who has been with me for all my student years and has supported and helped all the time.

# ABSTRAKT

Odhad věku dožití dospělých jedinců je stále velmi problematická součást odhadu biologického profilu jedince. Současné metody odhadu věku jsou schopné věk odhadovat přesně a spolehlivě pouze ve velmi širokých věkových rozmezích. Ke zlepšení této situace nepřispívá ani fakt, že jsou tyto postupy založeny většinou jen na užití jednoho kosterního indikátoru, ať jsou věkové změny hodnoceny vizuálně, či se opírají o nevhodné statistické přístupy. Disertační projekt vychází z těchto nedostatků a jeho cílem je přispět k lepšímu porozumění problematiky a především do výzkumu začlenit a otestovat nové technologie a výpočetní přístupy. Charakter disertační práce je spíše metodologický a celý výzkum byl rozdělen na tři oddělené části, čemuž odpovídá i stanovení tří hlavních cílů.

Prvním cílem bylo aplikovat v multi-populačním souboru vizuálně hodnocených dat (n=941) dvou artikulačních plošek pánevní kosti různé sofistikované matematické postupy s cílem dosáhnout přesnějšího odhadu věku dožití. Výsledné hodnoty RMSE, které byly v rozmezí 12 až 14 let však o zpřesnění odhadu nevypovídají.

Dále jsme se zaměřili na zhodnocení výstupů z různých skenovacích zařízení a vlivu případných odchylek v zachycení povrchu kloubní plochy pubické symfýzy na analýzy odhadu věku. Soubor 29 pánevních kostí české středověké populace byl naskenován dvěma skenery, HP 3D SLS a NextEngine. Jako referenční údaje byly zvoleny skeny povrchu s vysokým rozlišením v malém vzorku (n=5) digitalizovaném skenerem Redlux Profiler. Výsledky srovnání digitalizovaného povrchu pubické symfýzy ukázaly, že ačkoli se povrchy získané ze dvou porovnávaných skenerů lehce lišily, tyto odchylky neměly významný vliv na analýzu odhadu věku dožití.

Třetím výzkumným zaměřením byla validace kvantitativní metody odhadu věku navržené Stoyanovou a kolektivem (2015, 2017) v odlišných populacích než referenční soubor severoamerických jedinců. Celkem bylo podrobeno kvantitativní analýze odhadu věku 96 3D modelů kloubních plošek pubické symfýzy ze smíšeného souboru evropských populací a 79 modelů pocházejících z Thajské populace. Výsledná chyba odhadu věku v celém souboru, bez omezení věku, byla příliš velká (RMSE hodnoty v rozmezí 15 až 22 let) a nepřijatelná. Nicméně v souboru omezeném do 40 let věku byly výsledné hodnoty chybovosti odhadu zřetelně nižší (RMSE hodnoty v rozmezí 6 až 8 let).



**Klíčová slova:** forenzní antropologie, bioarcheologie, biologický profil, odhad věku dožití, virtuální antropologie, 3D skenování, populační specificita

# ABSTRACT

Age-at-death estimation of adults is still a very difficult part of estimation of individual's biological profile. Current age estimation methods can estimate age accurately and reliably only as very broad age ranges. The fact that these approaches are mostly based on only one skeletal indicator, whose age changes are assessed visually, or are based on an inappropriate statistical approach, does not contribute to the improvement of this situation. The dissertation project is based on these shortcomings and it aims to contribute to a better understanding of the issue and to integrate and test new technologies and computational approaches in research. Given the methodological nature of this dissertation, the whole research was divided into three separate parts, which corresponds to the establishment of three main objectives.

The first objective was to apply different sophisticated mathematical techniques in a multi-population database of visually evaluated data (n=941) of two articulation surfaces of *os coxae* to achieve a more accurate estimate of age. However, the resulting RMSE, which ranged between 12 to 14 years, do not indicate an accurate estimate.

Furthermore, we focused on the evaluation of outputs from various scanning devices and we tested the influence of possible differences in captured surfaces on the estimation of age. A sample of 29 pelvic bones of the Czech medieval population was scanned by two scanners, HP 3D SLS and NextEngine. In addition, a small sample (n=5) was digitized with a high-resolution Redlux Profiler scanner whose outputs were utilized as reference surfaces. A comparison of the digitized pubic symphyseal surfaces showed that, although the surfaces obtained from the two compared scanners varied slightly, this did not significantly affect the age analyses performed.

The third research focus was the validation of the quantitative age estimation method proposed by Stoyanova et al (2015, 2017) in different populations. A total of 96 3D models of surfaces of pubic symphysis from the European samples and 79 models from the Thai sample were subjected to a quantitative analysis of age estimation. The resulting error of age estimation across the whole dataset, without age restriction, was too large (RMSE ranged 15 to 22 years) and unacceptable. However, in the sub-sample under 40 years, the resulting values were significantly better (RMSE between 6–8 years).



**Key words:** forensic anthropology, bioarchaeology, biological profile, age-at-death estimation, virtual anthropology, 3D scanning, population specificity

# LIST OF ABBREVIATIONS AND EXPLANATIONS OF TERMS

| | |
|---|---|
| 3D | Three dimensional |
| AODE | Averaged One-Dependence Estimator |
| BMD | Bone mineral density |
| CEI/XXI | 21st Century Identified Skeletal Collection |
| CISC | Coimbra Identified Skeletal Collection |
| CT | Computer tomography |
| DF | Discriminant function |
| DISH | Diffuse Idiopathic Skeletal Hyperostosis |
| DNA | Deoxyribonucleic acid |
| DXA | Dual-energy x-ray absorptiometry |
| HP 3D SLS | HP 3D Structured Light Scanner PRO S2 |
| KKU | Khon Kaen University |
| M3 | Third molar (wisdom tooth) |
| MAE | Mean absolute error |
| NDE | No Dependence Estimator |
| OA | Osteoarthritis |
| OPD | Osteon population density |
| RMSE | Root mean square error |
| S1 | First sacral vertebra |
| S2 | Second sacral vertebra |
| SAH | Slice and Algee-Hewitt score (regression model of Stoyanova's method) |
| SIJ | Sacroiliac joint |
| TA | Transition analysis |
| TPS/BE | Thin plate spline/bending energy (regression model of Stoyanova's method) |
| TSP | Two-step procedure |
| VC | Ventral curvature (regression model of Stoyanova's method) |
| Accuracy | Success of the method to correctly identify the biological profile parameter |
| Precision | Accuracy of the measurement or scoring of the characters included in methods |
| Reliability | Dependability of the method in the process of its validation in another population |

# TABLE OF CONTENTS

# 1. INTRODUCTION

Age-at-death estimation of adults is one of the most important and problematic tasks when building a biological profile, which forensic anthropologists and bioarchaeologists have to deal with. Skeletal and dental development, which is very useful for subadult age estimation, is, however, complete in adults. Therefore, researchers can only rely on the physiological degeneration processes of these structures for their estimates.

The visual assessment of age-related changes observed on various skeletal indicators is by far the most widely used approach and its history dates back to at least 1920 (Todd, 1920). Since then, countless methods and their revisions have been developed based on various skeletal structures in an effort to estimate age-at-death. At the same time, doubts have arisen about the emerging methodologies and skeletal indicators used, which resulted from the inability to estimate the age of adults accurately and reliably. The limits originate, for instance, from the use of inappropriate statistical methods, from the misconception that one indicator can capture the whole age range of adult life, from the visual nature of the assessment of morphological changes, or from the inappropriate use of population-specific methods. More importantly, it is necessary to remember that variability in the aging process is a daily routine caused by internal and external factors.

Two major milestones have helped to address some of the shortcomings. The first one was a court case resulting in the so-called Daubert criteria, which were primarily directed to methods of the forensic sciences (Daubert v. Merrell Dow Pharmaceuticals, 1993; Grivas and Komar, 2008). These criteria require, for example, that methods be tested, be known among the scientific community and be sufficiently successful. The second very considerable milestone was the Rostock manifesto (Hoppa and Vaupel, 2002), which established theoretical recommendations for paleodemography in response to required biostatistical improvements. A very important output was the Transition analysis, which significantly helped to improve age estimation. The Daubert criteria and the Rostock manifesto further influenced the development of aging methods in both, paleodemography and forensic anthropology.

Very soon after medical imaging techniques and later scanning technologies became more accessible, they began to integrate into anthropological disciplines. They offer a new data source in the form of virtual models derived from computed tomography

or surface scanners, complementing osteological collections of universities, museums, and other institutions. They also provide access to new type of information and innovative ways to analyse it (e.g. (Verhoff et al., 2008; Grabherr et al., 2009; Ramsthaler et al., 2010)).

Even though many steps towards increasing the accuracy and reliability of age estimation methods have been undertaken, the age estimation of adults is still unsatisfactory. Accurate and reliable estimates are possible only as very broad age ranges. However, at least a partial solution lies in the involvement of imaging technologies and computational quantitative analyses of age-related surface changes. This PhD project aims to contribute to a better understanding of aging processes, and to implement new technologies (e.g. scanning technologies) along with the approaches they offer (e.g. analysis of surface changes), while respecting the Daubert requirements.

The first part (*Part I: Background*) of this present doctoral thesis is divided into three main chapters. The first chapter, *Biological profile indicators in forensic anthropology and bioarchaeology*, focuses on age-at-death estimation in the context of biological profile and briefly on the population specificity of biological profile methods. Furthermore, the chapter is devoted to the estimation of age in juveniles and young adults, with an emphasis on the accuracy of given skeletal indicators and methods. There is an introduction to the principles of the age estimation of adults, as well as the methods of adult age estimation at the microscopic level, i.e. biochemical and histological methods. The background of this thesis contains one of the author's own publications that focused on the issue of population specificity.

The second chapter of the first part, *Visual traits in adult age-at-death estimation: A change of paradigm*, is devoted to an overview of gross age indicators and the relevant and most commonly used visual methods, with focus on the three joints of the pelvic bones. Furthermore, the factors that can influence joint morphology and hence age estimation are summarized, as well as the main limitations to skeletal aging methods and new perspectives.

The third chapter, *Age-at-death estimation: New approaches and insights*, presents new approaches and methods of virtual anthropology that have entered the field of anthropology in recent years and its implementation into aging methods. The conclusion of

the chapter is dedicated to new quantitative approaches to the evaluation of age changes in joint surfaces and forms a transition bridge to the second part.

Important topics are summarized and discussed throughout the dissertation thesis, at the end of each chapter.

The research part of this dissertation thesis (*Part II: Personal contribution to age-at-death estimation methods*) is based on three of the author's own publications concerning aging methods and scanning devices (Kotěrová et al., 2018a; b, 2019). Three main objectives have been set. Firstly, we aimed to improve the age estimates obtained from visually assessed skeletal indicators through the application of several mathematical approaches. Secondly, the research was dedicated to the evaluation of outputs from various scanning devices, and to the question of whether the possible differences would be reflected in the analyses of age-at-death estimation. Thirdly, the project also aimed to validate the already published computational aging approach proposed by Stoyanova et al. (2015, 2017) in geographically diverse populations. Therefore, the chapters *Materials and Methods* and *Results and Discussion* are divided according to each research focus. The doctoral thesis is finally concluded with a direction for future research and an overall summary.

# Part I: Background

## 2. BIOLOGICAL PROFILE INDICATORS IN FORENSIC ANTHROPOLOGY AND BIOARCHAEOLOGY

***The "big four"***

Positive individual identification is of predominant interest to forensic experts and forensic anthropologists when only skeletal remains are left. The priority is to build a biological profile, which is the first step to successfully identifying unknown skeletal remains. The biological profile consists of four main osteological analyses (demographic characteristics), sometimes called the "big four" (Figure 1): assessment of sex, age, biogeographic origin (ancestry estimation, population affinity) and stature (e.g. (Cattaneo, 2007; Franklin, 2010; Hartnett et al., 2018; Ellingham and Adserias-Garriga, 2019)). Estimation of the biological profile is also an integral part of bioarchaeology and paleodemography (or demographic anthropology), and the methods applied overlap to a large extent with those used in forensic practise. However, this is not always the case, as the standards for identification are sometimes different because the main objective in bioarcheology is to reconstruct the demographic structure (Hoppa and Vaupel, 2002; Cunha et al., 2009; Godde and Hens, 2012). In the bioarchaeological context, the demographic profile and human lifeways, including health status and mortality, of past population are also reconstructed (DeWitte, 2017).

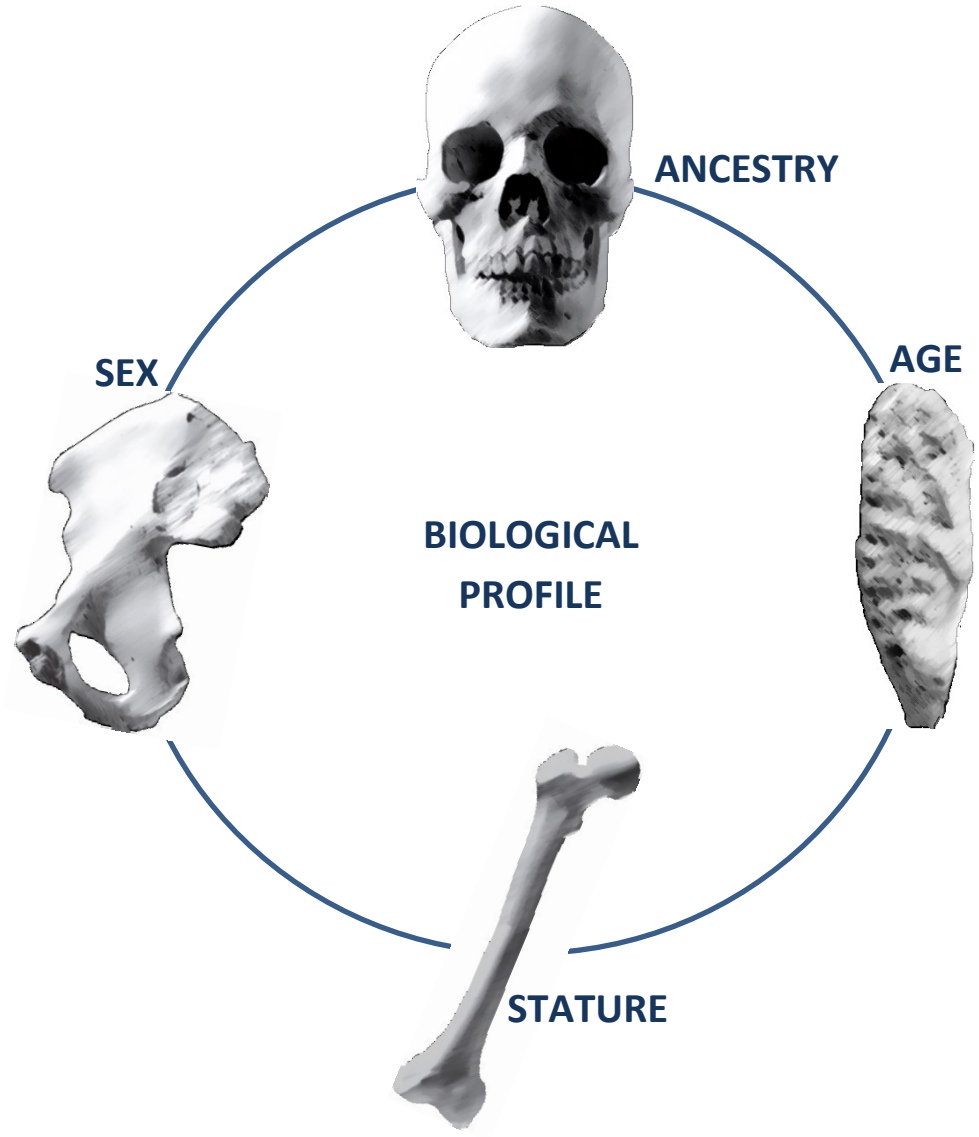


**Figure 1.** Parts (structures) of skeleton most commonly used to estimate biological profile parameters.

When the whole skeleton is available, sex estimation of adult skeletal remains is not problematic; however, this is not always the case. Very accurate estimates are made possible with morphological or metric methods based on *os coxae* (e.g. (Phenice, 1969; Brůžek, 2002; Murail et al., 2005; Brůžek et al., 2017)), which is bone presenting marked sexual dimorphism which is not burdened with population specificity (Garvin, 2012; Brůžek et al., 2017). This means that methods developed from the *os coxae* are independent of geographic location. The cranium also bears sexually dimorphic skeletal elements, but the level of dimorphism is lower than of *os coxae,* and its expression is population specific (L'Abbé et al., 2013). Assessment of sex that is based on the cranium can usually be done with 80–90% accuracy (İşcan and Steyn, 2013b). The sexual

dimorphism of the postcranial skeletal structures, which is generally small in humans, is related to body size and proportions. The postcranial elements can also provide very accurate (over 90%) sex estimates (Spradley and Jantz, 2011). However, it differs among populations and can be affected by a secular trend, which can have serious impact on the use of methods (Cabo et al., 2012; İşcan and Steyn, 2013b; Berg, 2017; Kotěrová et al., 2017). While with *os coxae* we can reach nearly 100% accuracy of sex estimation (Brůžek et al., 2017), the accuracy is slightly lower when other skeletal elements are employed, most commonly with the use of metric variables and discriminant function analysis. Moreover, the information about sex can be very helpful for other analyses.

Even though the knowledge of sex in juveniles is as important as in adults, the estimation is much more problematic, almost impossible. This is due to the fact that sexually dimorphic traits develop only under the influence of testosterone whose levels are low before puberty. Nevertheless, the ongoing research in this area, is still coming up with new studies, using both metric and geometric morphometry approaches, especially based on the pelvic bones or long bones (e.g. (Bilfeld et al., 2013; Luna et al., 2017; Stull et al., 2017)). At the moment, however, genetic diagnosis of sex is recommended as far as juvenile sex estimation is needed (Berg, 2017).

As was mentioned, sex estimation is population-specific (except for the *os coxae*-based approaches), thus the knowledge of an individual's biogeographic origin is a great advantage and a way to estimate sex accurately (Spradley and Weisensee, 2017). Undoubtedly, population affinity is the most controversial parameter of the biological profile (Spradley and Weisensee, 2017); it often encounters criticism for obvious reasons that may not necessarily be reported here. The cranium is most often employed for the estimation of ancestry, and the morphology of the dentition is also used. There are plenty of publications that could generally be divided into three categories, according to the approaches used to assess population affinity. Firstly, there are visual-based approaches (e.g. (Hefner, 2009; Hefner and Ousley, 2014; Hefner et al., 2014; Scott et al., 2018)), then craniometric data-based approaches (e.g. (Hefner et al., 2014; Navega et al., 2015a; Katherine Spradley and Jantz, 2016)) and finally shape-based analyses of 3D models (e.g. (Murphy and Garvin, 2018; Musilová et al., 2019)).

Similar to sex estimation, assessment of the geographical origin of juveniles is very difficult because the relevant traits are not yet fully developed (Nikita, 2017a); thus, the assessment is not often recommended (Christensen et al., 2014a). However, few studies

have been published and some authors have found cranial variation to be geographically dependent (population specific) from an early age (Wood, 2015).

However, it could be said that the community of researchers is divided into those who believe that individuals can be assigned to a particular population and those who are much more sceptical (İşcan and Steyn, 2013c). The results are rather unconvincing, and in medico-legal cases, DNA analyses for biogeographical ancestry should be preferred. The salient question is whether, in the future, globalization will not result in even less clear geographical differences, which would make it impossible to assess the geographical origin (İşcan and Steyn, 2013c).

Stature can be estimated based on two main approaches: the mathematical method and the anatomical method. The former uses regression formulae that are calculated based on correlations of individual skeletal elements, usually long bones, which are highly correlated to total stature (Raxter et al., 2006). This approach is beneficial when the human remains are incomplete (Raxter and Ruff, 2018). The anatomical method, on the other hand, summarizes the lengths and heights of all skeletal elements contributing to the total height, and adds soft tissue correction factors (Fully, 1956; Raxter et al., 2006), thus providing more accurate estimates. Given that stature changes with age even in adults (Friedlaender et al., 1977), adjustments for that are often provided, as well as equations that do not include age for cases where the age is not known (e.g. (Raxter et al., 2006; Niskanen et al., 2013)).

The stature estimation is not an exception and its assessment in juveniles is problematic, as in the case of sex and biogeographic origin. Among the main reasons for this is that the epiphyses are not fused yet and thus, are very often missing. Therefore, the length of long bone diaphysis is used for stature estimation (Smith, 2007; İşcan and Steyn, 2013d). Moreover, sex and biogeographic origin assessments are extremely difficult in subadults, which may, together with quite variable pattern of growth among individuals, increase the inaccuracy of stature estimates (İşcan and Steyn, 2013d; Nikita, 2017b).

The secular trend is inherently associated with stature estimation, which can cause considerable bias when using inappropriate regression equations derived from old reference samples (e.g. (Jantz Meadows and Jantz, 1999; Nikita, 2017b)).

Age-at-death assessment of adults is considered the most critical and yet the most challenging parameter of the biological profile, not only for forensic anthropology, but also for paleodemography (DeWitte, 2017). Apart from sex estimation, which narrows down

the range of potential positive matches by about 50%, the evaluation of age also significantly helps to narrow down the selection (Algee-Hewitt, 2017), thus reliable and accurate age estimates are of utmost importance. The knowledge of age also contributes to the creation of a demographic profile in the case of mass graves, or to distinguish commingled human remains (Adams and Byrd, 2014). Nevertheless, adult age assessment is based on the observation of degenerative changes on the joint surfaces, which are highly unpredictable and affected by many factors. Thus, age-at-death estimation is surrounded by great discussion and effort to find a reliable and accurate age estimation tool (Cattaneo, 2007). This is demonstrated by a study that analysed publications in the field of forensic anthropology from 2008–2017. According to their research, age estimation was the most common topic after sex estimation (Lei et al., 2019). The main subjects of debate are the age indicators themselves, the analytical techniques and statistical methods, as well as the usability among different recent populations and in archaeological populations (Ubelaker and Khosrowshahi, 2019).

Unlike other aspects of the biological profile, age estimation in juvenile individuals is the only parameter that could be estimated reliably and accurately, i.e. the estimates have low bias and are within an admissible range of error (İşcan and Steyn, 2013e). This is because aging relies on predictable bone and dentition growth and development.

### *Population specificity*

In the medico-legal and bioarchaeological contexts, *´population´* can be construed as a group of individuals that are associated to a particular geographical area. On the basis of some genetic similarity and on similar external influences, they exhibit a certain range of variability (Franklin and Flavel, 2019). Each population was exposed to different environmental and socio-economic influences and has a different genetic background, which results in variation in human development (Olze et al., 2004; Schmeling et al., 2006) and various expressions of sexual dimorphism (Ubelaker and DeGaglia, 2017).

Biological anthropologists have to work with this fact and should be increasingly aware of the importance of population-specific standards. The performance of methods originally developed on one population could in fact be diametrically opposed when applied to another population, resulting in unknown variation (Franklin and Flavel, 2019). This is true for all biological profile parameters, with the exception of biogeographic origin

assessment, where population variability in skeletal morphology is the core of these analyses. Probably the most discussed  is the topic of population specificity in relation to sex estimation based on non-pelvic skeletal elements (e.g. (Franklin et al., 2013; Ubelaker and DeGaglia, 2017; Hora and Sládek, 2018; Čechová et al., 2019)). Here, the variability of skeletal dimensions within populations and different degrees of sexual dimorphism of skeletal elements could cause misclassification and complete failure of sex estimates when population specificity is neglected (e.g. (Walker, 2008; Kotěrová et al., 2017; Franklin and Flavel, 2019)).

Kotěrová and colleagues (2017) empirically quantified the size of the error when population-specific discriminant functions derived from measurements of the tibia were applied in a Czech sample (see the *Appendix A.1*). Published classification functions for the Portuguese, south European (Spanish, Italian, and Greek) and North American populations were applied on a sample of 56 virtual models of the modern Czech population. Results clearly show that the discriminant functions developed in geographically and chronologically diverse populations and applied to a modern Czech population result in the complete failure of sex estimation. The results of comparisons of discriminant function performance in the original samples (for which the discriminant functions were designed) and in that of the Czech population are shown in Table 1.

**Table 1.** Application of DFA proposed in different populations in recent Czech population: simulation of disregarding DFA population specificity (Kotěrová et al., 2017).

| | Classification success in reference sample (%) | | | | Classification success in Czech sample (%) | | | |
|---|---|---|---|---|---|---|---|---|
| **Discriminant function** | **Male** | **Female** | **Pooled mean** | **Sex bias** | **Male** | **Female** | **Pooled mean** | **Sex bias** |
| **DF1 (Portuguese)[1]** | 84.8 | 83.7 | 84.2 | 1.1 | 90.0 | 38.5 | 66.1 | 51.5 |
| **DF2 (Portuguese)[1]** | 84.8 | 83.7 | 84.2 | 1.1 | 93.3 | 38.5 | 67.9 | 54.8 |
| **DF3 (Portuguese)[1]** | 82.6 | 87.8 | 85.3 | –5.2 | 83.3 | 53.9 | 69.7 | 29.4 |
| **DF4 (Portuguese)[1]** | 80.4 | 77.6 | 79.0 | 2.8 | 100.0 | 0.0 | 53.6 | 100 |
| **DF5 (Spanish)[2]** | 95.2 | 92.0 | 93.4 | 3.2 | 100.0 | 11.5 | 58.9 | 88.5 |
| **DF7 (Greek)[2]** | 89.4 | 85.9 | 87.8 | 3.5 | 100.0 | 3.9 | 55.4 | 96.1 |
| **DF8 (pooled)[2]** | 84.1 | 87.6 | 86.0 | –3.5 | 100.0 | 7.7 | 57.1 | 92.3 |
| **DF9(Euro-American)[3]** | 85.0 | 84.6 | 84.8 | 0.4 | 100.0 | 7.7 | 57.1 | 92.3 |

DF = discriminant function
[1] (Brůžek, 1995); [2] (Kranioti and Apostol, 2014); [3] (İşcan and Miller-Shaivitz, 1984)

Population specificity was also observed in the aging process (e.g. (Schmitt et al., 2002; Bassed et al., 2011; Buk et al., 2012)), so researchers are advised to be aware of methods developed from distant populations (Purves et al., 2011). This issue will be discussed further below (*Chapter 3.3.*).

Given that body proportions and height may vary among populations or regions, population (region) specific as well as sex specific standards are required for stature estimation (Langley, 2017).

If population affinity (biogeographical origin) is known before age, sex and stature are estimated, population standards should be used, as long as they are available for the particular population. Nowadays, the use of appropriate standards (contemporary and population-specific), which lead to accurate biological profile estimation, is generally respected, accepted and recommended (e.g. (Franklin et al., 2013, 2015; Ubelaker and DeGaglia, 2017)). However, we should consider the cases, which are common, in which population specificity is unknown. In recent years, the rate of immigration has markedly increased, resulting from ever-evolving globalization, free borders among the European states, human trafficking and political conflicts around the world, often leading to the mass movement of people across borders (e.g. (Obertová and Cattaneo, 2018; Schaefer et al., 2018)). All together, they contribute to irreversible change in the composition of

populations. This is becoming increasingly topical and urgent. Thus, when human skeletal remains are found and their population affinity is unknown, it is uncertain whether they actually belonged to the particular area (Brůžek and Murail, 2006; Spradley et al., 2008; L´Abbé and Steyn, 2012). Then the question is whether the population-specific methods of biological profile can provide relevant information; it is now known that it definitely does not necessarily do so.

***Conclusion***

Apart from biogeographical origin estimation, to which there are conflicting views, the greatest challenge and at the same time very important part of the biological profile is age-at-death estimation (e.g. (Ritz-Timme et al., 2000; Cunha et al., 2009)), even though a positive identification through this parameter alone is not possible. The need to find both accurate and reliable approaches can be substantiated by dozens of new publications each year that are trying to propose new methods based on various skeletal age indicators. So far, we have lacked such methods, despite the fact that researchers have been dealing with skeletal age-at-death estimation for at least one hundred years (e.g. publication of (Todd, 1920)). Age estimation in living individuals is not the subject of this thesis; however, the use of some methods overlaps due to imaging technologies (radiography, computed tomography).

The success rate of age estimation decreases with increasing age. We are able to estimate age in juveniles relatively accurately and reliably, but after the completion of growth and development, accuracy and reliability decrease rapidly (e.g. (Cunha et al., 2009; İşcan and Steyn, 2013e)). The subject of this thesis is adult age-at-death estimation; however, to compare the age intervals and accuracy with which we operate in juveniles and vice versa in adult individuals, we first must look at age estimation of subadult skeletal remains (*Chapter 2.1.*). For the period of early adulthood, when the individual is an adult legally (this threshold may vary slightly among countries), but the skeletal development is still ongoing, the fusion of the medial end of the clavicle epiphysis, dental mineralization or fusion of several apophyses can be used (e.g. (Langley et al., 2017; Hartnett et al., 2018)). This will be the subject of *Chapter 2.2*. A brief introduction into the senescence processes of the adult skeleton (*Chapter 2.3.*) and an overview of some marginally used biochemical and histological methods (*Chapter 2.4*.) will close this section.

## 2.1. Juvenile age estimation based on the skeleton

First of all, it is crucial to stress here that biological age (whether skeletal or dental) is the one that can be estimated in physical anthropology, although it is the chronological age (actual number of years passed from birth to death) that is wanted (Cunha et al., 2009; Christensen et al., 2014b). Skeletal age estimation is based on the correlation between biological and chronological age, which, however, is never perfect, nor can be, because skeletal aging is variable and influenceable. The relationship between chronological and biological age is closer in younger individuals (Acsádi and Nemeskéri, 1970; Nawrocki, 2010; Christensen et al., 2019).

The known timing of skeletal changes across the lifespan and their correlation with chronological age both enable researchers to estimate age with variable accuracy. In juveniles the entire skeleton changes as a function of growth, and the timing of developmental processes is very predictable (Franklin, 2010; Langley et al., 2017; Nikita, 2017c). The process of development and growth from the immature to mature skeleton includes the appearance of ossification centres, epiphyseal and apophyseal union and development, and the eruption of teeth (Ubelaker and Khosrowshahi, 2019). The development of the subadult skeleton may be influenced primarily by internal factors (such as genetic); however, during the developmental phase of an individual's life, the external factors (e.g. illness, malnutrition) may also have an effect. In contrast, degenerative changes in adulthood are predominantly influenced by external factors (Christensen et al., 2014b).

Recognizing whether human remains belonged to a juvenile individual or an adult usually is not a problem for an anthropologist. Childhood is a part of the lifespan when age estimation is the most accurate, because the correlation with chronological age is the closest. However, these changes with age and the gap between the timing of the changes and the chronological age widen in adulthood (Nawrocki, 2010). In fact, knowledge of juvenile age is an important prerequisite for the possible estimation of other aspects of the biological profile. At the same time, age estimation is the only parameter of biological profile that can be estimated with accuracy when identifying juvenile skeletal remains (Nikita, 2017a; Cunningham, 2019).

Both, dental development (root development, mineralization and eruption of both, deciduous and permanent teeth) as well as skeletal growth and development (degree of

ossification, long bone length and epiphyseal union) can be used to estimate the age of juveniles (e.g. (Ritz-Timme et al., 2000; Baccino et al., 2013; İşcan and Steyn, 2013f; Langley et al., 2017)). However, dental development is very strongly correlated with chronological age, thus is the best choice for estimating age, since it offers more accurate and reliable estimates than bone maturation (e.g. (Cunningham et al., 2016; Langley et al., 2017; Ubelaker, 2018a; Adserias-Garriga, 2019; Ubelaker and Khosrowshahi, 2019)). This is because tooth development is less affected by external factors and is under stronger genetic influence (Thesleff and Sharpe, 1997; Merwin and Harris, 1998; Nikita, 2017c). On the contrary, skeletal maturation is often affected by environmental influences (Cunningham et al., 2016; Langley et al., 2017) and is less affected by genetic factors. Given these facts, a bad state of health, (e.g. illness, malnutrition) is more likely to affect long bone growth than tooth calcification, which can thus result in a greater split between dental and skeletal age (Langley and Tersigni-Tarrant, 2017).

The current review of the most commonly used methods for estimating the age-at-death of juvenile individuals is provided by recent forensic manuals (İşcan and Steyn, 2013a; Christensen et al., 2014c; Langley and Tersigni-Tarrant, 2017; Nikita, 2017d; Latham et al., 2018; Adserias-Garriga, 2019). These methods, which are common and are considered important by the author of this thesis, will be chosen and overviewed.

The subadult period is usually divided into several subgroups (age categories), but the age intervals differ slightly in different forensics manuals and publications. Such classes may be fetus, newborn, infant (0–6 years), child (7–12 years) and adolescent (13–20 years) (Baccino et al., 2013). The choice of the appropriate aging method depends, of course, on the skeletal remains available for age analysis and on the assumed subadult age category (Ritz-Timme et al., 2000; Baccino et al., 2013). In the prenatal stage (foetuses) and the early postnatal period, age can be reported within weeks or months. Later, with increasing age, estimates are possible in a range of 1–3 years (İşcan and Steyn, 2013e); Reppien et al. state that age estimates for small children can be within the range of 2 years, and for subadults within the range of 4 years (Reppien et al., 2006).

### 2.1.1. Methods based on dental development

The dental age can be assessed according to tooth mineralisation (tooth formation) or the timing of tooth eruption. However, methods based on mineralisation are considered to be more accurate in contrast to methods based on eruption (Solheim and Vonen, 2006;

Langley et al., 2017). Tooth eruption is influenced by both, external factors such as insufficient or unbalanced food intake (Gatta et al., 2008; Gaur and Kumar, 2012), infection, pathology, and trauma (AlQahtani, 2019), and internal, such as the hormones of a relatively unstable thyroid gland (Loevy et al., 1987; Chandna and Bathla, 2011). Methods for evaluating stages of dental mineralisation are much more reliable because they are not so affected by external factors and are under the control of hypophysis (Langley et al., 2017).

Teeth start to develop around the 6$^{th}$ week *in utero* (Caruso et al., 2016; Cunningham et al., 2016) and continue into early adulthood, making them a very useful indicator covering the entire juvenile period (Caruso et al., 2016; Ubelaker, 2018a). Among the most commonly used dental aging methods for subadults is the evaluation of the developmental stages of each available tooth (dental crown and root mineralization) with the use of relevant illustrations. Such methods include the method of Moorrees et al., based on tooth mineralization and eruption of permanent and deciduous teeth (Moorrees et al., 1963a; b), and that of Demirjian et al. (Demirjian et al., 1973; Demirjian and Goldstein, 1976). Moorrees considered both mandibular (Moorrees et al., 1963a; b) and maxillary teeth (Moorrees et al., 1963a). This method can be used from the 4$^{th}$ month *in utero* to 25 years of age, while Demirjian based his approach on mandibular dentition for 2.5–17 year old children. These methods are sex-specific. Several modifications and adaptations for a variety of populations have been published (e.g. (Smith, 1991; Willems et al., 2001; Chaillet and Willems, 2004; Liversidge and Molleson, 2004; Blenkin and Evans, 2010)).

Some very user-friendly methods based on dental development are probably the ones developed as an atlas or charts (Nikita, 2017c) that depict the stages of dental development at different ages (composite visual system). A well-known and most commonly used chart was authored by Ubelaker (Ubelaker, 1978, 1989), who modified the chart of Schour and Massler (Schour and Massler, 1941), detailing the deciduous and permanent tooth. Ubelaker's chart includes tooth eruption (gingival), as well as tooth formation for individuals aged 5 months *in utero* to 35 years (Figure 2). Another atlas, known as the London Atlas of Human Tooth Development, was recently proposed by AlQahtani et al. (AlQahtani et al., 2010), and is available online at www.atlas.dentistry.qmul.ac.uk. The atlas consists of stages between 30 weeks *in utero* and 23.5 years. Both tooth formation and eruption (alveolar) are used.

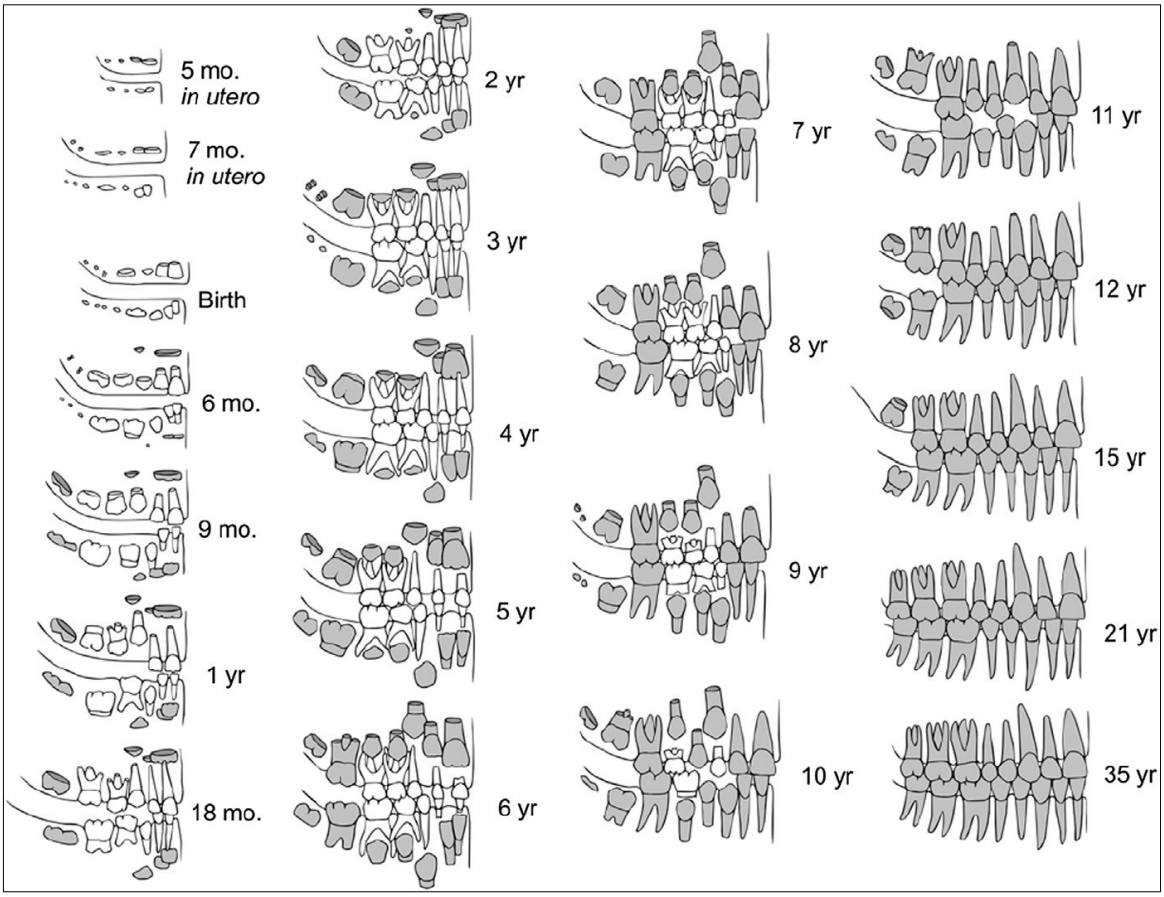


**Figure 2.** Deciduous and permanent tooth development and eruption; Chart published by Ubelaker (1989) from (Christensen et al., 2014c).

There are a few serious limitations to all the methods assessing dental development. First of all, as a result of secular changes, the methods derived from modern samples should not be applied to archaeological material, as the applicability is uncertain (Nikita, 2017c). Secondly, since the estimation of sex in juvenile skeletal remains is almost impossible and females generally mature more quickly than males, the estimation of age when sex is unknown can be affected (Molinari et al., 2004). Another problem is the population specificity of the timing of dental formation, although it is not as pronounced as it may be in skeletal development (Ubelaker, 2018a). The application of methods to other populations is limited and generally the use of a population specific standard of dental development is recommended (Willems et al., 2001; Langley et al., 2017; Nikita, 2017c).

Dentition is invaluable for estimating age in juveniles, especially in foetuses, newborns and infants (Franklin, 2010; Baccino et al., 2013; Langley et al., 2017). Nevertheless, around 14–15 years of age, most teeth are usually fully developed and erupted, except for the third molars, which remains the only teeth undergoing

morphological development even later, thus being useable for age estimation (Reppien et al., 2006; İşcan and Steyn, 2013f; Draft and Kasper, 2019). However, their development tends to be more variable in contrast to the tooth formation during infancy, which is the least variable (AlQahtani et al., 2010).

### 2.1.2. Skeletal age assessment of subadults

When dentition is not available, skeletal development and maturation (i.e. diaphyseal length of long bones, the appearance and union of primary ossification centres and the union of epiphyses) are very good tools to estimate age as well (İşcan and Steyn, 2013e). The correlation of diaphyseal length with chronological age is especially useful in foetuses, where the relationship is very strong and the intersexual and interpopulation differences are small. This is because the development is not yet as influenced by extrinsic factors (Fazekas and Kósa, 1978; Hoffman, 1979; Franklin, 2010). Although it dates back to the late 1970s, Fazekas and Kósa's standards are still being largely used by some researchers (Fazekas and Kósa, 1978). However, doubts about the suitability of these standards in modern times are justified due to secular trend. More recent sources (e.g. (Schaefer et al., 2009; Cunningham et al., 2016)) and many population-specific studies in literature are available. Regression equations to estimate fetal age exist for those from 3 to 9 months (Fazekas and Kósa, 1978). As children grow, the body height (or stature) becomes increasingly variable among them, and measurement of bones becomes less helpful (İşcan and Steyn, 2013d).

The appearance and fusing of epiphyses is a regular consistent process with a known timing sequence that correlates with chronological age (Langley et al., 2017). Some widely used reference textbooks include the one by Cunningham et al. (Cunningham et al., 2016) and Schaefer et al. (Schaefer et al., 2009), where the original work of Fazekas and Kósa (1978) as well as modern standards are all included. Most epiphyses finish their union between 10 and 25 years which makes them useful even for age estimation in young adults.

Similarly, as it was already mentioned for dental development, stages of maturation are the same for all individuals, but the rate of particular developmental milestones may be affected by many factors (Liversidge et al., 2015). The development could differ between sexes and populations, as well as groups of people with different socio-economic or geographical background. If possible, population specific standards should be used (Nikita,

2017c; Schaefer et al., 2018; Ubelaker and Khosrowshahi, 2019). Skeletal maturation could, of course, be influenced by nutritional and health state (Prentice et al., 2006).

The next chapter will be devoted to late-fusing skeletal areas, i.e. parts of the skeleton that still interfere with the late adolescent and young adult periods.

## 2.2. Skeletal maturation and aging in late adolescents and young adults

Late adolescents (16–20) and young adults (20–30 years) can be identified as individuals in which skeletal maturation is being completed (epiphyseal union and finishing of third molar development). This period is sometimes called the transition phase. Secondary ossification centres fuse with the primary mostly during adolescence, but a few of them fuse later than most others (Baccino et al., 2013; Cunningham et al., 2016). Given that, these late-fusing bone markers hold the potential to be used as age indicators of early adulthood. However, the literature is often inconsistent with regard to the recommendation of skeletal indicators to determine the onset of adulthood (reviewed in (Falys and Lewis, 2011)). Population variability has already been emphasized several times and is again necessary to bear in mind.

### *Crista iliaca*

One of the late-fusing anatomical structures is the *spina iliaca,* which starts its fusion around 14–18 years in female and 17–20 years in males, and completes fusion around 20–23 years of age (Schaefer et al., 2009). The approximate appearance and fusion times of ossification centres on *os coxae* are shown in Figure 3.

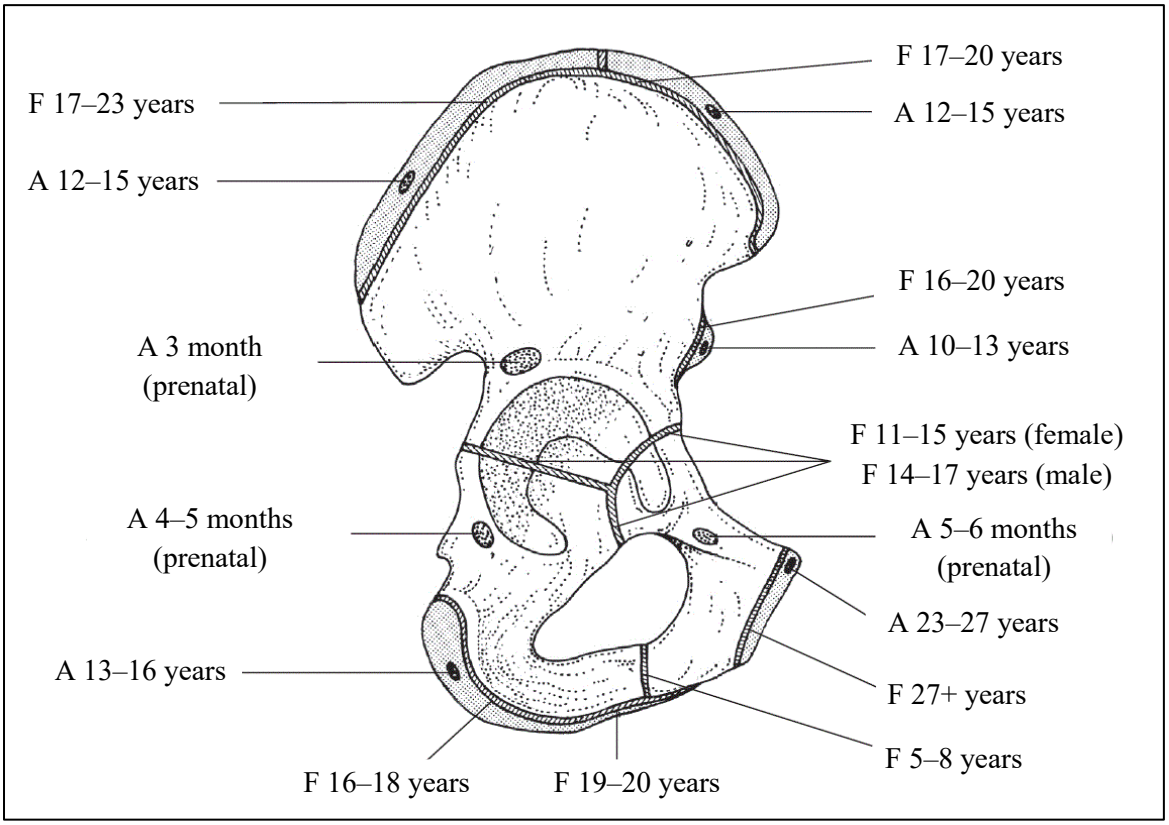


**Figure 3.** Summarized times of the appearance and fusion of pelvic ossification centres.

In the mid-twentieth century, the so-called Risser sign grading system was developed with the original intention to evaluate iliac crest maturation (ossification) to provide additional information for patients with scoliosis (e.g. (Risser, 1958)). Later, apart from clinical purposes, the Risser sign began to be used for forensic age estimation, especially in living individuals (e.g. (Wittschieber et al., 2013a; Lottering et al., 2017; Bartolini et al., 2018)). Besides the Risser sign grading system, there are other methods, for instance, the Kreitner and Kellinghaus method (Kreitner et al., 1998; Kellinghaus et al., 2010) originally developed for clavicle ossification and modified by Wittschieber et al. for application on the iliac crest (Wittschieber et al., 2013b).

### *Spheno-occipital synchondrosis*

Another marker is the union of the spheno-occipital synchondrosis (Figure 4), which is typically considered a transition point from juvenile age to adulthood

(Cunningham et al., 2016). It was initially reported that this synchondrosis fuses at around 18 years of age, between the ages of 17 and 20–25 (Nikita, 2017c; Hisham et al., 2018).

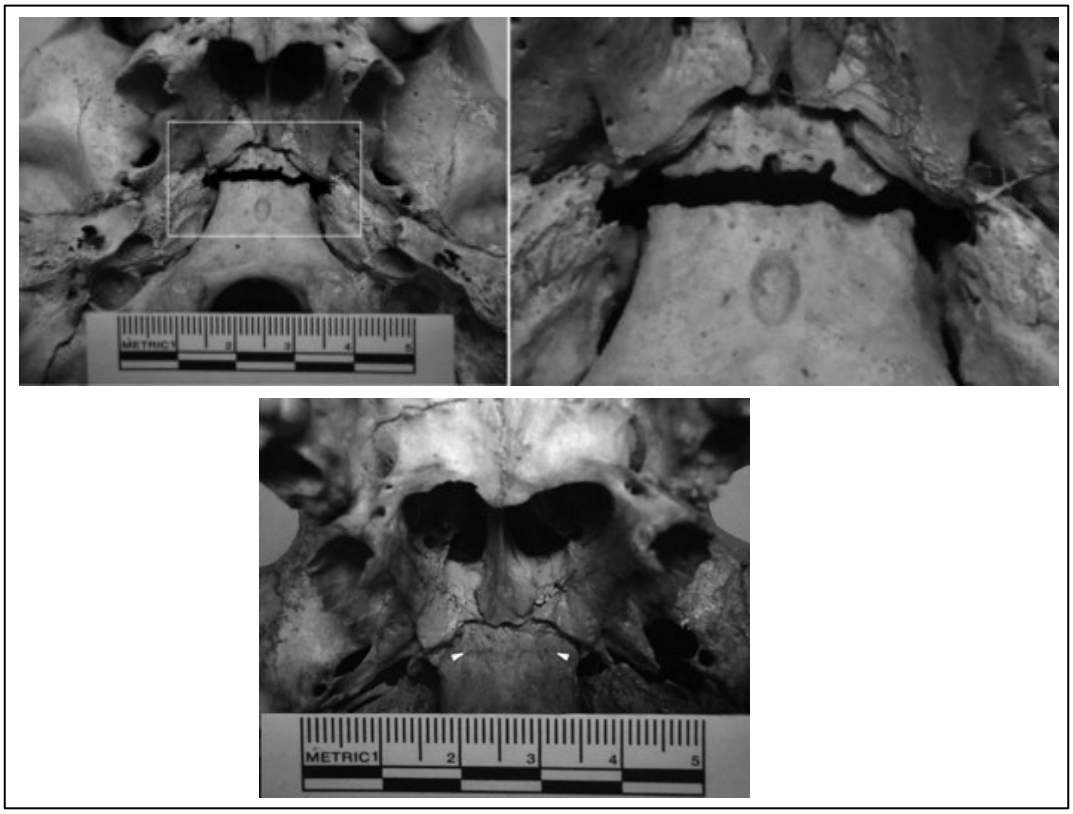


**Figure 4.** Unfused spheno-occipital synchondrosis (top left and right), fusing synchondrosis (bottom). Taken from (Shirley and Jantz, 2011).

However, it is now known that the fusion can occur much earlier in juvenile age, approximately from 11 years in females and from 13 years in males, but the timing is widely variable between populations. However, the average age of the union remains around 19 and 20 years, with slightly earlier occurrence in females (e.g. (Shirley and Jantz, 2011; Franklin and Flavel, 2014; Cunningham et al., 2016; Hisham et al., 2018)).

### *Third molar*

When second molars are erupted, third molars are the ones that are still developing (İşcan and Steyn, 2013f; Draft and Kasper, 2019). Crowns start to form between 6–12 years (Cunningham et al., 2016) and do not erupt earlier than the late teens (around 18 years) to the early twenties (Christensen et al., 2014b; Langley et al., 2017). This is also

related to the completion of root formation, which may occur even after 25 years (Adserias-Garriga, 2019) and is often advanced in males (Mincer et al., 1993; Willershausen et al., 2001). The crown and root development of the third molar can be rated, for instance, by the Mincer age estimation technique (Mincer et al., 1993). The development of the third molar is reflected also in the London Atlas of Human Tooth Development (AlQahtani, 2019) and Ubelaker's sequence diagram (Ubelaker, 1978, 1989).

Unfortunately, M3 is not only characterized by the most variable timing of development among all the teeth, which makes it not recommended as a sole age indicator, but also by its frequent congenital absence (Mincer et al., 1993; Rozkovcová et al., 2005; Christensen et al., 2014b; Algee-Hewitt, 2017; Ubelaker, 2018a; Draft and Kasper, 2019). Nowadays, the third molar is quite often extracted (e.g. (Marciani, 2007)) which excludes it as an age indicator.

### *Thoracic vertebrae, Sacrum*

The complete fusion of the annular epiphysis (which is usually the last one to fuse) of the thoracic vertebrae starts at around 17–18 years till the late twenties (Schaefer et al., 2009; Cardoso and Ríos, 2011).

Generally, the secondary ossification centres of the sacrum fuse later than the thoracic, cervical or lumbar vertebrae. In the sacrum, first (S1) and second (S2) sacral vertebrae are the last to fuse. The complete union is seen at around 25 years or more (McKern and Stewart, 1957; Cunningham et al., 2016); in fact, it could be observed as late as at 35 years of age (Belcastro et al., 2008).

### *Clavicle*

The last epiphysis to fuse is usually the medial end of the clavicle (Figure 5). The fusion of the ossification centres used to be assessed via computed tomography (Kellinghaus et al., 2010; Bassed et al., 2011; Ekizoglu et al., 2014; Gurses et al., 2016), X-rays (Schmeling et al., 2004; Garamendi et al., 2011), magnetic resonance imagining (Hillewig et al., 2013; Schmidt et al., 2017) and last but not least, via dry bones (Langley-Shirley and Jantz, 2010).

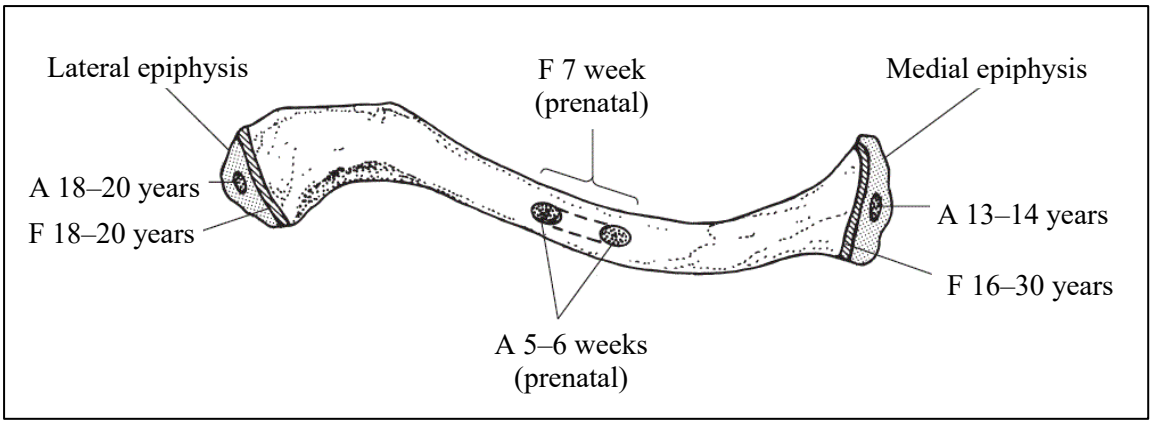


**Figure 5.** Times of appearance and fusion of ossification centres of the clavicle. A = appearance, F = fusion. Modified (Cunningham et al., 2016).

A summary of several publications concerning the timing of epiphyseal fusion is presented in Table 2. Generally, there are several scoring systems used to record the stage of epiphyseal fusion of the clavicle. Probably the most often used scoring system is the 5-stage system developed by Schmeling and colleagues (2004). The five phases are as follows: (I) the ossification centre has not yet ossified; (II) the ossification centre has ossified, but the epiphyseal cartilage has not ossified; (III) the epiphyseal cartilage is partially ossified; (IV) the epiphyseal cartilage is fully ossified; (V) the epiphyseal cartilage has fused completely and the epiphyseal scar is no longer visible. Another system is the one of McKern and Stewart (McKern and Stewart, 1957). Here five stages are defined as (1) no union, (2) beginning union, (3) active union, (4) recent union, and (5) complete union. As can be seen from Table 2, there is considerable variation in fusing time. The minimal age when complete fusion was observed ranged from 17 to 26.4 years in males and from 20 to 26.7 years in females. Such differences may be due to various input data, discrepancies in technical parameters (e.g. slice thickness) or simply due to population variation. If this is the case, the use of population-specific standards is crucial here. At the same time, it has to be pointed out that there is some overlap between the years when the epiphysis may still not be fused in one individual, while it may be completely fused in another (e.g. (Schaefer et al., 2018)). The influence of a secular trend in the fusing time in the clavicle was observed (Langley-Shirley and Jantz, 2010).

**Table 2.** Comparison of epiphyseal fusion times (stage IV and V of Schmeling et al. and McKern and Stewart scoring systems) in the clavicle.

| Publication | Study population | Scoring system | Age of sample (years) | Min-max age at stage IV (years) | | Minimal age - complete fusion (years) | |
|---|---|---|---|---|---|---|---|
| | | | | Male | Female | Male | Female |
| **Bassed et al. (2011)** | Australia | (1) | 15–25 | 17–25* | 19–25* | 17 | 20 |
| **Ekizoglu et al. (2015)** | Turkey | (1) | 10–35 | 20–35* | 20–34 | 25 | 25 |
| **Garamendi et al. (2011)** | Spain | (1) | 5–75 | 19.7–33 | | 20.6 | |
| **Gurses et al. (2016)** | Turkey | (1) | 10–35 | 21.0–35.9* | 20.9–35.9* | 25.0 | |
| **Hillewig et al. (2013)** | Belgium | (1) | 16–26 | 22.1–26.9* | 18.1–26.9* | not stated | |
| **Kellinghaus et al. (2010)** | Germany | (1) | 10–35 | 21.6–35.8* | 21.3–35.2* | 26.4 | 26.1 |
| **Langley-Shirley et al. (2010)** | USA | (2) | 11–33 | 22–32 | 20–31 | 19 | 24 |
| **Schmeling et al. (2004)** | Germany | (1) | 16–30 | 21 | 20 | 26.0 | 26.7 |
| **Schmidt et al. (2017)** | Germany | (1) | 12–24 | 21.7–24.8 | 21.3–24.9 | not stated | |

(1) Schmeling et al., 2004; (2) McKern and Stewart, 1957
* Maximal age in sample

Once all epiphyses and apophyses are completely merged and dentition is fully developed and erupted, in other words, skeletal and dental maturity is reached, the individual may be considered fully grown from an anthropological point of view. While late-fusing skeletal structures complete their fusion, other anatomical structures begin to degenerate, and techniques relying on degenerative processes of the skeleton are the ones to be used to estimate age-at-death (Baccino et al., 2013; Adserias-Garriga and Wilson-Taylor, 2019; Ubelaker and Khosrowshahi, 2019). The evaluation of the degenerative

processes of the joints by its very nature allows the estimation of age only with wide age intervals. Adult age estimation will be the main content of the next part of this thesis.

## 2.3. Principles of adult age-at-death assessment

The age-at-death estimation of juveniles is often both accurate and reliable; however, for adults this is not the case, especially for older individuals. While subadult age estimation techniques rely on skeletal growth and development, post-developmental senescence and degenerative changes of various skeletal and dental structures are most often the basis for estimating age-at-death once the skeleton is fully matured (Cunha et al., 2009; Baccino et al., 2013). These physiological degeneration changes have relatively predictable order, but the rate of their occurrence varies very much depending on the genetic background and external influences of an individual which are responsible for the separation of biological and chronological age (Gocha et al., 2015). Because of that, the biology of aging is very variable among populations. The effect of external factors on age-at-death assessment will be the subject of *Chapter 3.2*. Visual methods based on assessing the gross morphology of the joint surface are the most widely and commonly employed to estimate age-at-death, because they are cheap, do not require bone or tooth destruction, are easy to apply and usually do not require special laboratory equipment.

However, organisms age at all levels. Age-related changes can therefore be observed not only on macroscopically observable external structures, but also on macroscopic internal structures, as well as at the microscopic or even biochemical levels.

The following chapter will provide a brief overview of methods that are not based on the visual macroscopic examination of the intact skeletal indicator, where some kind of intervention into bone or dental structure is needed. Traditional macroscopic age-at-death estimation methods will be the content of *Chapter 3,* with a more detailed focus on pelvic bone indicators and methods.

## 2.4. Biochemical and histological methods for adult age estimation: Principles and their accuracy

### *Biochemical methods of age estimation*

The aging of the organism is accompanied by natural changes that happen at different levels. In this chapter, changes at the lowest level – biochemical – will be briefly overviewed. When it comes to chemical modifications occurring throughout life, all possible approaches are based on the modifications of protein configuration leading to alterations in cells and tissues. Of all the chemical approaches, the racemization of aspartic acid is the most accurate and reliable (Zapico et al., 2019a). The essence of racemization is converting L-amino acids to D-amino acids with age. It is a natural process that every metabolically stable protein undergoes (Masters et al., 1977; Arany and Ohtani, 2010). The racemization of aspartic acid is one of the fastest racemization, which is of great benefit in forensic sciences. To estimate age, the changing ratio of L and D form is used (Zapico and Ubelaker, 2013). Among the tissues suitable for analysis (i.e. tissues with slow metabolic rate), dentin appears to be the more accurate and the most used for age estimation, as it produces strong correlation with age and provides a standard estimation error of ± 1.5–4 years (Ritz-Timme et al., 2000; Ohtani and Yamamoto, 2010; Adserias-Garriga and Zapico, 2018). Moreover, dental tissues are durable, therefore, the dentition is often preserved. In fact, aspartic acid racemization is one of the most reliable methods of age estimation based on dentition (Zinni and Crowley, 2017). On the other hand, in the case of bone tissue, the use of this is difficult (Ohtani, 1998). This chemical approach can be performed if teeth are present, if skeletal remains were not burnt and the time elapsed since death does not exceed several decades (Ritz-Timme et al., 2000). Nonetheless, it has its limitations; for instance, the lack of expertise in laboratories, which is required due to complex procedures, or the dependence on tooth type (Minegishi et al., 2019), as well as its destructive nature. It has been reported that this approach provides more accurate estimates of life expectancy than methods based on bone or dental morphology (Ritz-Timme et al., 2000).

Other chemical methods that were studied in order to estimate age are lead accumulation, collagen crosslinks, chemical composition of teeth and advanced glycation

end products (Zapico and Ubelaker, 2013; Zapico et al., 2019a), although they do not reach such accuracy.

Apart from chemical approaches used to estimate age, there are also molecular biology approaches, which use various modifications to the DNA. They are, however, not as accurate as aspartic acid racemization (Zapico et al., 2019b). Among the studied approaches, there are mitochondrial DNA mutation, signal joint T-cell receptor excision circle rearrangement, epigenetic modification and telomere shortening (Zapico et al., 2019b). Telomere shortening research is well known among the molecular methods and offers about 7 years' error (Tsuji et al., 2002). Very recently, research has begun to focus on age-related epigenetic modification, specifically DNA methylation (Bocklandt et al., 2011; Bekaert et al., 2015; Giuliani et al., 2016). Initial studies reported an error of about 5 years with the use of saliva or blood samples (e.g. (Bocklandt et al., 2011; Zbieć-Piekarska et al., 2015)); however, the more recent ones reached very high accuracy. For example, Giuliani et al. predicted the age from dental pulp with error (median of the absolute difference) of 2.25 years; from tooth cementum with error of 2.45 years; but from cementum, with error of 7.07 years (Giuliani et al., 2016). DNA methylation is becoming a very promising molecular marker of age for use in forensic sciences; however, more research is required.

### *Bone and dental histological aging*

Histomorphometric ageing methods take advantage of the fact that the bone remodels itself over the span of a lifetime from primary lamellar bone, through primary osteons to final secondary osteons (Christensen et al., 2014b). The proportion of remodelled bone increases with age. The approaches are based, for instance, on counting the numbers of different structures or their density, which in the case of intact and fragmentary osteons is referred to as osteon population density (OPD). The density grows with age until the maximum OPD is achieved (Crowder et al., 2012). The microscopic changes can also be expressed as the percentage of remodelled or unremodelled bone, using metrical or area variables. However, the remodelling rates are bone-specific (Robling and Stout, 2007; İşcan and Steyn, 2013e; Christensen et al., 2014b). In addition to OPD, the size and shape of osteons was also assessed, and it was found that while size decreases with aging, osteon circularity and OPD both increase (Goliath et al., 2016). However, these

optimistic results were not confirmed by a validation study (Lagacé et al., 2019). The most often employed skeletal elements for histological analysis include the femur, ribs and tibia (İşcan and Steyn, 2013e).

The histomorphometric approaches are useful, especially in fragmentary human remains, when macroscopic indicators are unavailable and morphological methods cannot be applied. However, even these approaches are burdened with a certain degree of subjectivity and there are various intrinsic and extrinsic factors influencing the remodelling rate (Robling and Stout, 2007; Christensen et al., 2014b; Crowder et al., 2018). Moreover, they are, of course, invasive and destructive, which is often undesirable. Other limiting factors that prevent these methods from being used are the equipment and training requirements. Bone structural biology has recently been imagined with the use of synchrotron radiation micro-CT (Maggiano et al., 2016; Crowder et al., 2018).

Just as the bone structure undergoes microscopic changes, so does the dental structure. The criteria that could be examined based on dentition are attrition, secondary dentin deposition, root transparency, root resorption, periodontosis and cementum apposition – the thickness of the band of cementum (Gustafson, 1950). Secondary dentin deposition, cementum apposition and root resorption can only be seen microscopically (Gustafson, 1950).

A special case of age estimation in humans is cementochronology, which has been used for many years to estimate the age of non-human mammals, but has been neglected for a very long time in humans. However, it has recently been on the rise (e.g. (Renz and Radlanski, 2006; Colard et al., 2015; Couoh, 2017; Bertrand et al., 2019a; b)) in both the forensic and bioarchaeological context (e.g. (Colard et al., 2015; Lanteri et al., 2018)). The annual apposition of cementum to the tooth root, which is the basis of cementochronology, shows very promising results in adult age estimation. The cementum deposits are characterized by dark and light bands that are counted (each pair = 1 year) at a microstructural level and added to the mean age of tooth eruption or, alternatively, root completion (Bertrand et al., 2019a). This apposition is characterized by a close relationship with chronological age and is applicable to individuals of all ages with permanent teeth (Ritz-Timme et al., 2000; Bertrand et al., 2019b). Today, counting cementum annuli is considered a reliable approach, providing strong correlation with age (Wittwer-Backofen et al., 2004; Bertrand et al., 2019a), although some exceptions exist (e.g. (Roksandic et al., 2009; Kasetty et al., 2010)). This may be due to the non-uniform protocols used in sample

preparation and/or the use of archaeological material where the applicability of method is not fully explored. The standardized procedure should be followed to reduce bias during preparation (see (Colard et al., 2015; Bertrand et al., 2019a)). The stated accuracy of the method is between ±2 to ±6 years (Grosskopf and McGlynn, 2011), which is better than the accuracy provided by traditional methods. However, the accuracy and reliability decrease with age, as the readability of incremental lines is more difficult with growing age (Bertrand et al., 2019a). The closer relationship between chronological and biological age is disturbed by intra and inter population variability in tooth eruption (Couoh, 2017). The mechanism of the annual increase has not yet been illuminated (Bertrand et al., 2019a), as is the case of the effect of periodontal diseases, as studies give contradictory conclusions (e.g. (Wittwer-Backofen et al., 2004; Dias et al., 2010)).

Nevertheless, the main disadvantage is the destructive nature of the sample preparation procedure, which requires the transverse section of the root (Figure 6). This limitation could be solved by non-destructive imaging using synchrotron micro-CT (Le Cabec et al., 2019), which so far has attained low accuracy but sufficient precision; further research is still needed.

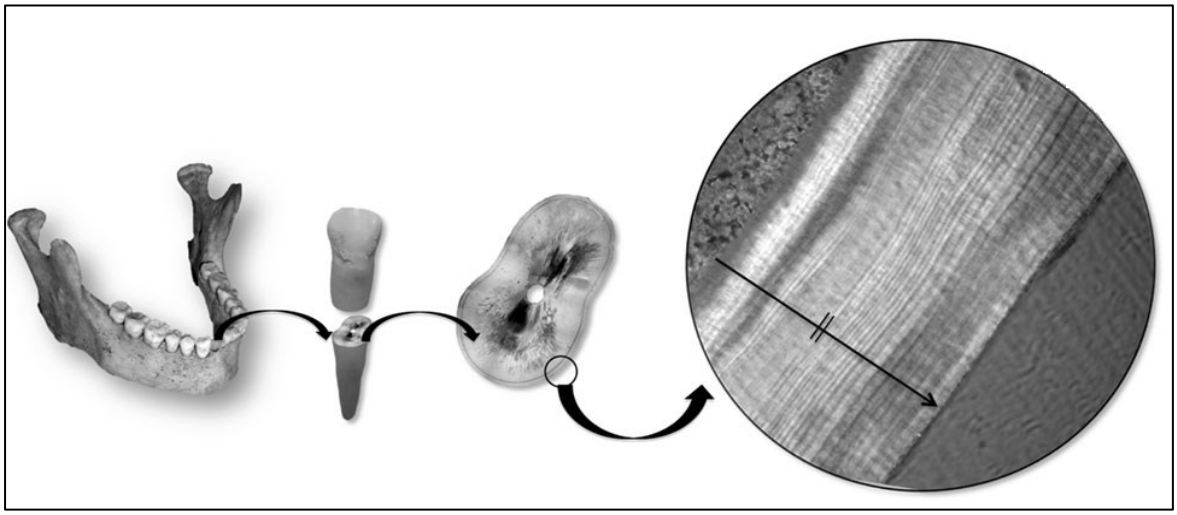

**Figure 6.** Principle of sample preparation in cementochronology: one root tooth, cross section, incremental lines in tooth root cementum. Modified (Colard et al., 2015).

***Conclusion***

Generally, the most reliable age estimation methods based on the dentition of adult individuals are the counting of cementum annuli, aspartic acid racemization on the enamel, dentin, or cementum (Rösing et al., 2007; Zinni and Crowley, 2017) and, recently, the epigenetic modifications (Giuliani et al., 2016). The best methods should not only be accurate, user-friendly, cheap and time-saving, but they should also be tested on different populations. This is often not the case with biochemical and histological methods, which is their frequent drawback. Moreover, another limitation is their need for laboratory equipment, proper training and experience, and the fact that they are in most cases destructive (Cunha et al., 2009; Lewis and Kasper, 2018). Given these reasons, conventional visual methods of age-at-death estimation are usually preferred, even though they often do not attain such high accuracy.

# 3. VISUAL TRAITS IN ADULT AGE-AT-DEATH ESTIMATION: A CHANGE OF PARADIGM

Approaches for adult age estimation are divided into several groups according to the level of the degeneration processes (micro and macroscopic) and their type, the structures available for examination and the like. The previous chapter briefly summarized some techniques based on bone microstructure, biochemical changes and tooth micro-modifications. Even though some of them could be very accurate and reliable, indisputably the most commonly applied methods in bioarcheology and forensic anthropology, besides the destructive dental methods, are the ones examining gross morphology; the metamorphosis of articular surfaces is the most common.

The choice of suitable skeletal indicators to estimate age-at-death is governed by several recommendations and requirements. They have to (1) show gradual changes with advancing age that are unidirectional and are highly correlated with chronological age. (2) The aging pattern should be reliably classifiable and measurable (according to the type of data). (3) The observable changes should happen approximately at the same time (consistent rate of change) in all individuals, or at least in the particular sex, biogeographic origin, etc. (4) Last but not least, the indicator should be as resistant to taphonomy as possible (Milner and Boldsen, 2012a; Algee-Hewitt, 2017). This ideal notion of how the indicators should look, however, clashes with reality in its biological variability, which makes it very difficult to find such an age indicator.

Traditional methods of age-at-death estimation rely on two different scoring systems: a phase-based or a component-based (composite) system. Both systems evaluate morphological features using regular patterns of degenerative changes, but the phase-based system describes an overall morphological change (i.e. more characters are grouped and assessed at the same time) and creates broad phases. On the contrary, component-based approaches first divide the particular age indicator into several components (traits) that are then scored independently. The latter is believed to be more objective and less dependent on observers experiences (Shirley and Ramirez Montes, 2015).

Since the end of the 20th century, the development of methods, especially in the forensic area, has been influenced by the requirements arising from the case Daubert v. Merrell Dow Pharmaceuticals, Inc. (1993), known as the Daubert criteria (Daubert v. Merrell Dow Pharmaceuticals, 1993; Grivas and Komar, 2008). The set of standardized

criteria target scientific expert witness testimony and must be met in order for the evidence to be admissible in court. For forensic science, including anthropological expertise such as age-at-death estimation, this means that methods must be presented to the scientific community through peer review and publication in renowned journals, and accepted by them. The methods and therefore the conclusions made by scientists must be testable and tested. The accuracy of methods must be known; the results should ideally be expressed via probability and the method should be sufficiently successful to be accepted (Daubert v. Merrell Dow Pharmaceuticals, 1993; Ritz-Timme et al., 2000; Grivas and Komar, 2008; Ousley and Hollinger, 2012; Langley and Tersigni-Tarrant, 2017). The Daubert standards changed the development of biological profile methods significantly.

The *Chapter 3* is dedicated to the morphological age indicators that are most commonly used for adult age estimation and for an overview of well-known and often employed traditional methods (*Chapter 3.1.*). Further, the influence of various factors on joint morphology used for age estimation will be debated (*Chapter 3.2.*). The last part of this section is devoted to the limitations that burden age estimation and possible solutions (*Chapter 3.3.*).

## 3.1. Age indicators and methods of age estimation

Contrary to growth and developmental changes of dentition and skeleton that are used to estimate the age of juveniles, indicators capturing postdevelopmental degenerative changes are used in the case of adults. Several articles reviewed age-estimation techniques that are the most commonly employed in a particular region, in particular anthropological literature (renewed journals) or/and the most used in an examined time period. Falys and Lewis (2011) examined aging techniques from three journals of the time period from 2004 to 2009. They reported that among the most frequently used osteological age indicators were dental attrition, cranial sutures, the pubic symphysis and the auricular surface of the ilium. They also stated that the use of population-specific standards were absent or very rare (Falys and Lewis, 2011). Another review was based on a questionnaire done by members of the American Academy of Forensic Sciences – Physical Anthropology section (Garvin and Passalacqua, 2012). The authors reached similar results; the most preferred skeletal indicators were the pubic symphysis, the sternal ribs end and the auricular surface. Conversely, cranial sutures and dental wear are among the least preferred.

The most common visual osteological aging methods are summarized and overviewed in anthropological textbooks. These are the most recent ones (Dirkmaat, 2012; İşcan and Steyn, 2013a; Christensen et al., 2014c; Langley and Tersigni-Tarrant, 2017; Nikita, 2017d; Latham et al., 2018; Adserias-Garriga, 2019). This thesis does not aim to be one of them; therefore, the list of methods and their revisions will not be complete. The attention will be directed to the sternal ends of ribs, which correlate with chronological age quite closely, to cranial sutures, which were widely used in the past and are sometimes used to this day (unfortunately), and shortly to dentition. However, this overview is focused mainly on the three articular surfaces of *os coxae,* the three age indicators that proved to be very useful, aiming to place their value in relation to different periods of adult age.

### *Cranial sutures*

Cranial sutures as an indicator of age-at-death went through periods in which they were widely used to estimate age, to a period when it was found that their obliteration had poor correlation with age, after which most practitioners stopped using these structures. It was observed that ectocranial, endocranial and also maxillary sutures gradually obliterate, and that these processes correlated somehow with age (e.g. (Todd and Lyon, 1924, 1925; Acsádi and Nemeskéri, 1970; Mann et al., 1991)). Among the first to be concerned with cranial sutures in relation to age and who suggested new appropriate methods were Todd and Lyon, who found that the sutures on the endocranium are the first to obliterate (Todd and Lyon, 1924, 1925). A widely used method was, for instance, the revised method of Meindl and Lovejoy, which is based on the component scoring of ectocranial sutures (Meindl and Lovejoy, 1985). However, the authors themselves concluded that suture closure should not be relied upon as a single indicator, but in association with other indicators (Todd and Lyon, 1924; Acsádi and Nemeskéri, 1970; Meindl and Lovejoy, 1985), or with the use of wide age intervals (Acsádi and Nemeskéri, 1970). Plenty of revisions and validations in different population samples have been carried out and have had very similar conclusions (e.g. (Powers, 1962; Galera et al., 1998; Donato et al., 2016; Ruengdit et al., 2018)). Some researchers agree with Masset's statement (Masset, 1990) that the relationship between actual age and the obliteration of cranial sutures is only statistical (Baccino and Schmitt, 2006). Even though complete obliteration is somehow

indicative of advanced age (Christensen et al., 2014b), it is not the rule, and using of suture closure should be avoided (Cunha et al., 2009; Baccino et al., 2013) or used only when a cranium without tooth is recovered; at least it should not be used alone, but as a general indicator in association with other indicators. Generally, suture closure is not among the preferred age-at-death indicators (Garvin and Passalacqua, 2012).

### *Sternal ends of ribs*

Sternal ends of ribs are elongated by costal cartilage of hyaline type (Standring, 2016), which received attention as an age indicator in the past (McCormick, 1980). Costal cartilage ossification is also associated with changes on the costal face (the area of the costochondral junction), which has become the subject of interest of several researchers.

İşcan and collective were the first who correlated morphological changes at the sternal end of the fourth rib with age and developed an aging method (İşcan et al., 1984a, 1985). This method was originally proposed for the fourth right rib as a phase-based system consisting of nine phases. The sternal end is initially flat or covered with billows, and its rim is rounded; later, the rim becomes irregular and sharp due to osteophyte formation, and also thinner with developing porosity (see Figure 7 for illustration). A series of casts (by France Casting) were created for users to facilitate the evaluation. Later, the use of the İşcan method was extended to other ribs of both sides, in addition to the fourth, since they go through similar changes (e.g. (Dudar, 1993).

**Figure 7.** Examples of morphological age-related changes of sternal rib ends. Individual 27 years (upper row), 49 years (middle row), and 85 years (lower row). Modified (Nikita, 2017c).

Few researchers focused on the metamorphosis of the first rib, and developed a method solely based on this structure (Kunos et al., 1999; Digangi et al., 2009). The original method was, of course, subjected to many revisions and adjustments, in order to attain better accuracy and make a validation in different skeletal samples (e.g. (Oettle and Steyn, 2000; Hartnett, 2010a)). However, probably due to variation among the skeletal samples and the not always ideal inter-observer repeatability the results are rather contradictory.

Nevertheless, rib techniques unlike, for instance, techniques based on pubic symphysis (see below) offer a longer period of metamorphosis, on the basis of which age can still be estimated accurately. On the contrary, the main disadvantage of ribs is their poor preservation ((Falys and Lewis, 2011; Cappella et al., 2017)).

### *Dentition*

Dentition provides very accurate and reliable age estimates in juveniles, and carries information about age changes even into adult age (as mentioned in the *Chapter 2.4.*). Several methods have been developed using teeth for age estimation. The best known methods include the Gustafson method (Gustafson, 1950), which has been modified many times, with the most significant revision by Lamendin (Lamendin, 1988; Lamendin et al., 1992). Gustafson incorporates six criteria that require a tooth section of which three must be evaluated microscopically (*Chapter 2.4.*). Among these criteria, attrition is the one that is the most affected by external factors; for example, by dietary factors. Lamendin's technique reduced the number of evaluated features to only two (root transparency and periodontosis). Moreover, the tooth section is not required, nor is there a need for special equipment or training. However, his approach is not recommended for individuals younger than 40 years (Lamendin et al., 1992), as the translucency is not present before 25 years of age (Baccino et al., 2014).

### *Pubic symphysis*

Pubic symphysis (Figure 8) is an anatomical structure by which scientists have undeniably been the most fascinated as an age indicator for a hundred years (for example the publication of (Todd, 1920)). A thorough systematic review of adult pubic symphysis was published by Becker and colleagues (Becker et al., 2010). Modern literature states that pubic symphysis is a secondary cartilaginous joint (Standring, 2016), although opinions have been different in the past (Becker et al., 2010). The two *os coxae* meet in the midline, with a fibrocartilaginous interpubic disc between them. In area where the disc is in contact with bone articulating surfaces, the surface of the pubic bone is covered by hyaline cartilage (Čihák, 2011; Standring, 2016). This joint is characterized by its very limited degree of movement; however, during parturition, mobility increases (Alicioglu et al., 2008; Mahato, 2016).

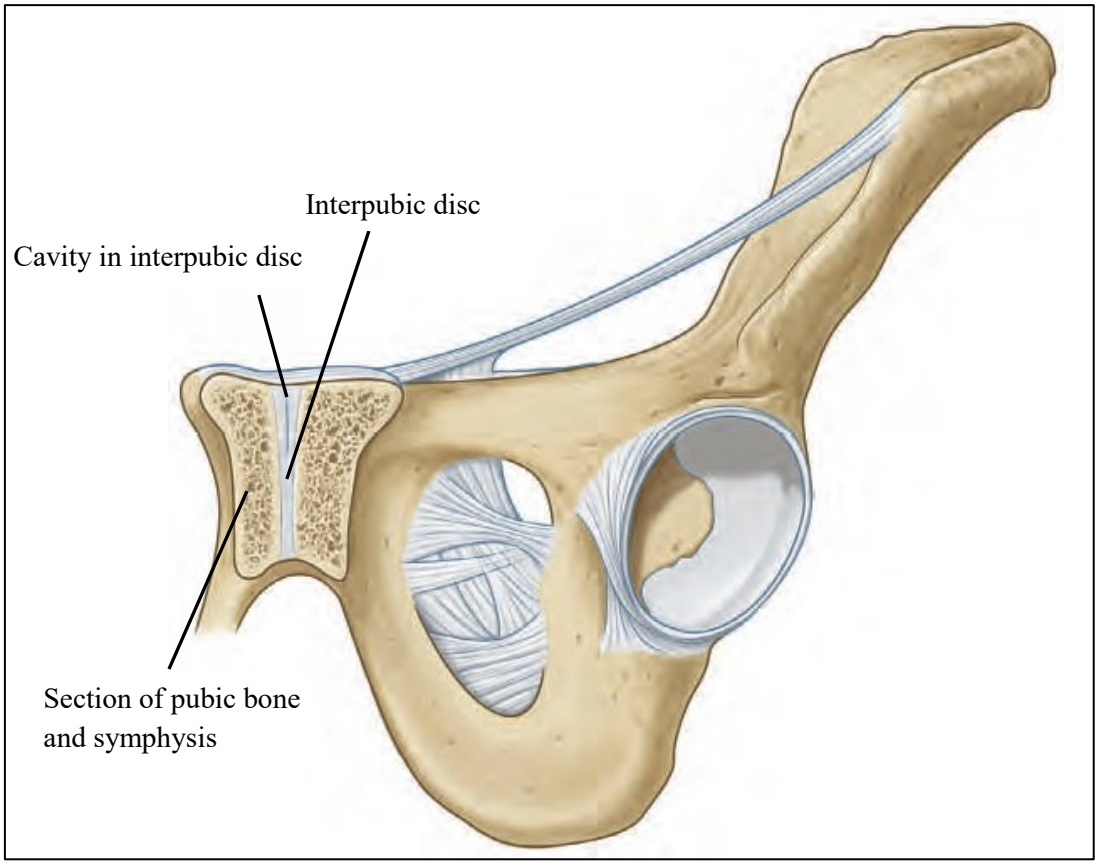


**Figure 8.** Coronal section of the pubic symphysis, anterior view of left bony pelvis. Modified (Standring, 2016).

The articulating surfaces undergo changes throughout adulthood, which have a degenerative character. The pubic face is at the beginning characterized by deep transverse ridges and furrows that begin to fade with age and are flat at the end. The texture of the bone also changes from smooth to fine and later to coarsely granular. Another feature subjected to age changes is the formation and subsequent deterioration (breakdown) of the rim. The formation of the rim starts on the dorsal margin, continuing to the ventral margin. Also during the aging process, porosity, pitting and overall erosion begin to appear. The changes were first observed and described by Todd as early as 1920, when he also designed a phase-based method (Todd, 1920). Todd's system was based on ten morphological phases starting from 18 years and ending with an open-ended interval of 50+ years (Todd, 1920; Milner and Boldsen, 2012b). Later, he suggested modifications of his system concerning the clustering of phases (1921). Since his publications, there have been countless tests, validations, modifications and new approaches developed, whether phase-based or component-based (e.g. (McKern and Stewart, 1957; Gilbert and McKern,

1973; Hanihara and Suzuki, 1978; Meindl et al., 1985; Brooks and Suchey, 1990; Lovejoy et al., 1995; Schmitt et al., 2002; Dudzik and Langley, 2015; Stoyanova et al., 2015; Savall et al., 2016; Merritt, 2018a)). The pubic symphyseal surface is most likely the most widely used indicator for estimating the age of adults (e.g. (Meindl et al., 1985; Garvin and Passalacqua, 2012; İşcan and Steyn, 2013e)). An important and most commonly used method is a modification of Todd's system published by Brooks and Suchey (Katz and Suchey, 1986; Brooks and Suchey, 1990). The symphyseal surface changes are in their approach visually scored into six phases and developed for each sex separately. Castings were made to facilitate the evaluation. For each phase, two bones presenting an early and a later pattern of the phase were chosen (Figure 9). Even the Suchey-Brooks method has received many revisions and validations on various population samples, including adjustment to the phases and statistical methods, both on dry bones (e.g. (Berg, 2008; Kimmerle et al., 2008; Hartnett, 2010b; Rivera-Sandoval et al., 2018)), and with the implementation of imaging and scanning technologies in anthropology, on virtual models (e.g. (Biwasaka et al., 2013; Lottering et al., 2013; Wink, 2014; Merritt, 2018a)).


**Figure 9.** Cast set for pubic symphyseal changes in female made by France Casting (http://www.francecasts.com/) for the Suchey-Brooks six phase system. The upper row corresponds to the early stage of phase; the lower row to the later stage (Langley et al., 2017).

However, already in his first publication, Todd stated this about the symphysis pubis: *"Indeed the symphysis tells its tale throughout life, although less clearly from forty years onward than at an earlier age"* (Todd, 1920). This conclusion seems to have been overlooked for a long time since many researchers applied the methods based on the evaluation of senescence changes in the pubic symphysis for the whole age range of adult life, which resulted in age estimates that were neither accurate nor reliable, particularly in individuals older than 40 years. The changes are simply too variable in older ages, as evidenced by the wide and overlapping age intervals published, for example, by Suchey-Brooks (Table 3).

**Table 3.** Descriptive statistics related to the Suchey-Brooks pubic age determination system (Brooks and Suchey, 1990).

| Phase | Female (n=273) | | | Male (n=739) | | |
|---|---|---|---|---|---|---|
| | **Mean** | **S.D.** | **95% range** | **Mean** | **S.D.** | **95% range** |
| **I** | 19.4 | 2.6 | 15–24 | 18.5 | 2.1 | 15–23 |
| **II** | 25.0 | 4.9 | 19–40 | 23.4 | 3.6 | 19–34 |
| **III** | 30.7 | 8.1 | 21–53 | 28.7 | 6.5 | 21–46 |
| **IV** | 38.2 | 10.9 | 26–70 | 35.2 | 9.4 | 23–57 |
| **V** | 48.1 | 14.6 | 25–83 | 45.6 | 10.4 | 27–66 |
| **VI** | 60.0 | 12.4 | 42–87 | 61.2 | 12.2 | 34–86 |

S.D. = standard deviation; all numerical values in years.

Nowadays, researchers are finally aware of and have stopped ignoring that a single age indicator cannot cover the whole age range; in the case of the symphysis pubis, its ability to reflect age is limited to approximately 40 years of life (e.g. (Buk et al., 2012; Baccino et al., 2014; Dudzik and Langley, 2015; Márquez-Grant, 2015)). For example, Dudzik and Langley (2015) proposed a component (composite)-based scoring approach for adults up to 40 years, which provides decision-tree-style flow charts. They offer narrower age ranges than the approaches assessing the overall morphological changes (Dudzik and Langley, 2015). Another approach that respects the biological ability of the pubic symphysis to show age-related changes is the Two-step procedure (TSP) of Baccino and co-workers (Baccino et al., 1991, 2014; Baccino and Zerilli, 1997), who combined two different skeletal structures – two different aging methods. The scoring systems of the Suchey and Brooks method and the Lamendin method are combined chronologically. The first three phases of the Suchey and Brooks method approximately cover adult age up to 40 years, while the Lamendin dental method is more appropriate for individuals older than 40 years (Lamendin, 1988; Lamendin et al., 1992). If the pubic symphysis is not available, the use of sternal end of ribs is recommended (Cunha et al., 2009; Baccino et al., 2014).

### *Auricular surface of the ilium*

The sacral and iliac auricular surfaces together create the sacroiliac joint (SIJ), which is currently considered to be partly synovial and partly syndesmotic (Mahato, 2016; Standring, 2016), as well as symphyseal (Puhakka et al., 2004). The type of joint was

debated a lot in the past (summarized in (Mahato, 2016)). Both the sacral and the iliac surfaces are covered with hyaline cartilage, which is thinner on the iliac surface (Puhakka et al., 2004; Standring, 2016) and is more fibrous and covered with fibrocartilage, contrary to the sacral part of the joint (Puhakka et al., 2004). Movements of SIJ are very limited, but increase during pregnancy (Standring, 2016).

The articular surface of the posterior pelvis (Figure 10) is characterized by metamorphic changes (remodelling of the joint surface) correlating with age; however, these were used to create age-at-death estimation methods much later than with the pubic symphysis. Activity in the retroauricular area can be registered (Lovejoy et al., 1985b).

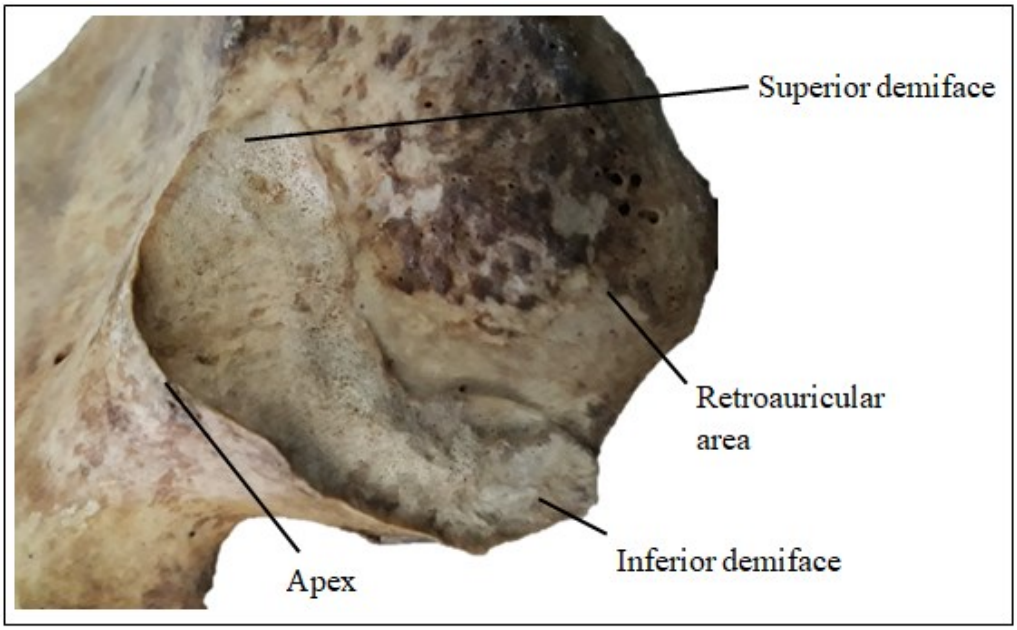


**Figure 10.** Posterior ilium – key areas of age-related surface changes used in methods of adult age-at-death assessment. Photo by Kotěrová A.

The advantage of the auricular surface over the pubic symphysis is that this area preserved much better (Lovejoy et al., 1985b; Osborne et al., 2004; Nikita, 2017c). Initially, the surface is regularly billowed (transversally organized); the billows later change to striae and eventually the surface becomes flat. Finely grained texture is gradually replaced by coarsely grained until the surface becomes dense at the end. The apex also undergoes degenerative changes, from a smooth and rounded surface to an irregular border with osteophytic lipping. With advancing age, micro and later macroporosity can be observed. Representative changes are depicted in Figure 11.

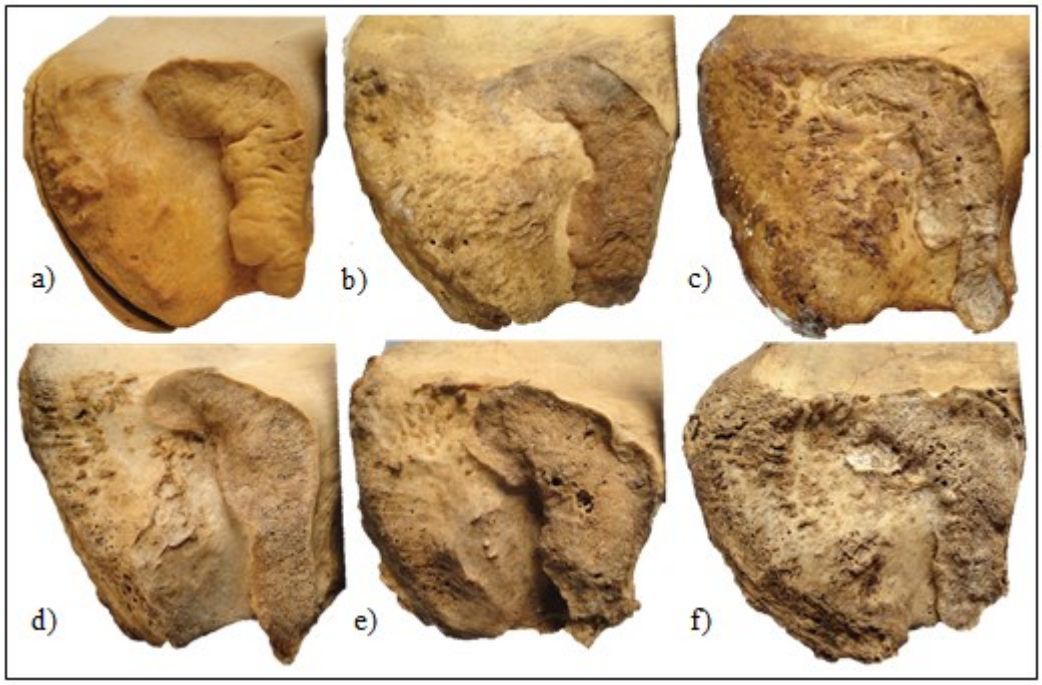


**Figure 11.** Degenerative changes (a–f) on the example of the left auricular surfaces and the retroauricular areas. Modified (Langley et al., 2017b).

Lovejoy and colleagues were the first to suggest a phase evaluation of the auricular surface and retroauricular area (Lovejoy et al., 1985b). They divided surface metamorphosis into eight phases at five-year intervals, where the last phase ends with an open interval of 60+ years. This original method was subjected to many validation studies and revisions (e.g. (Buckberry and Chamberlain, 2002; Osborne et al., 2004; Schmitt, 2004)) which found out that the age of older individuals is underestimated and that the 5-year intervals are too narrow, making age estimates inaccurate. Osborne et al. reported only 33% accuracy with the 5-year age intervals (Osborne et al., 2004). Buckberry and Chamberlain revised the Lovejoy et al. method by proposing a composite scoring system

based on five characteristics, excluding the retroauricular area which they did not find accurate (Buckberry and Chamberlain, 2002). Nevertheless, the age ranges that correspond to the seven stages assigned according to the composite score are unacceptably wide. Moreover, some of the descriptions are difficult to follow (Márquez-Grant, 2015). Studies that followed their approach reported rather inconsistent results (e.g. (Mulhern and Jones, 2005; Falys et al., 2006)).

### *Acetabulum*

The hip joint is a synovial ball-and-socket joint. The articular surface of the acetabulum is called the lunate surface, which is covered by articular cartilage. The non-articular area, the acetabular fossa (Figure 12), is filled with fibroelastic fat covered by synovial membrane. The fibrocartilaginous rim that is attached to the acetabular margin is called the acetabular labrum (Standring, 2016; Ward, 2016). Unlike the other joints on the *os coxae*, the mobility of the hip joint is considerable (Standring, 2016).

The surface of the acetabulum is relatively protected (Calce and Rogers, 2011), thus its preservation is better, than for instance the preservation of the pubic symphysis.

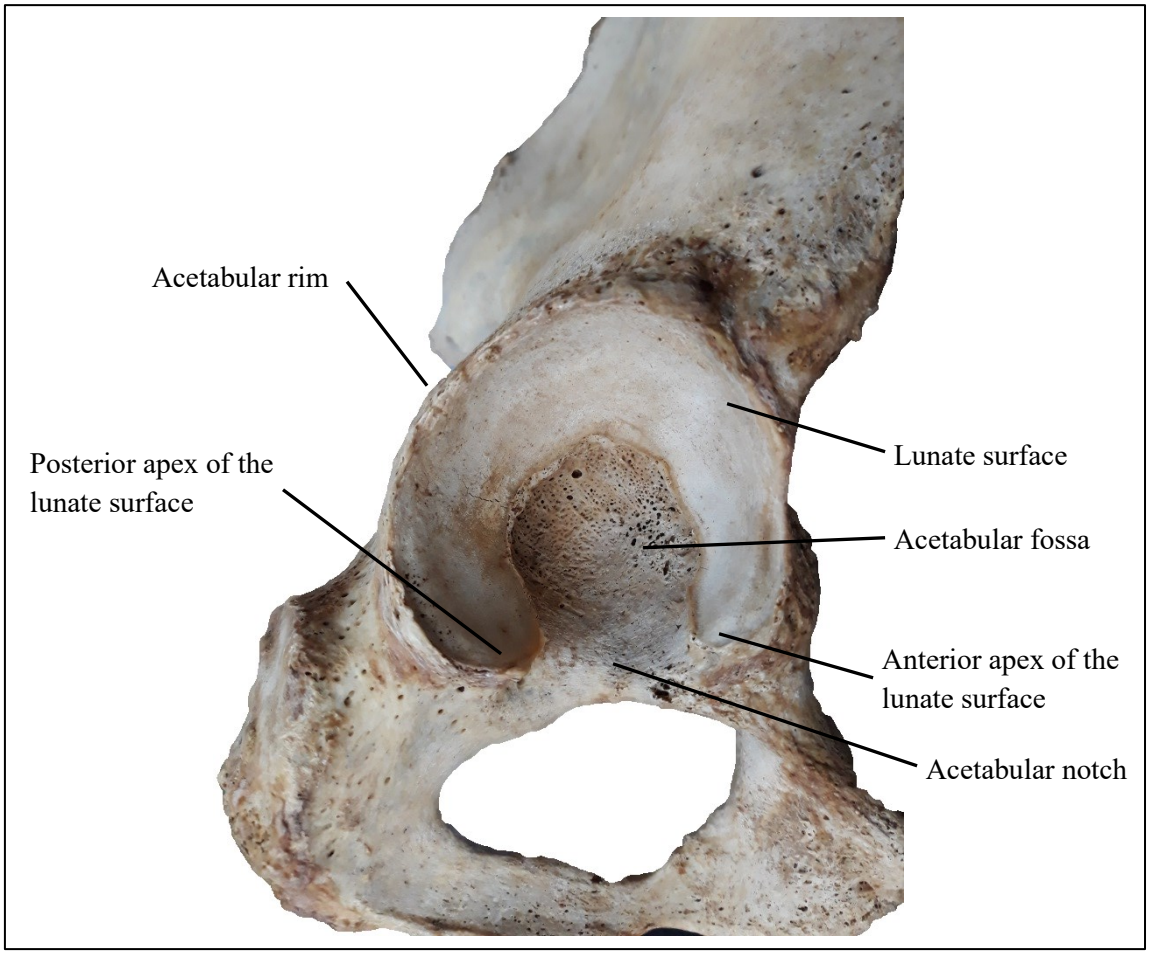


**Figure 12.** Acetabulum - key areas of age-related surface changes used in methods of adult age-at-death assessment on the example of right *os coxae*. Photo by Kotěrová A.

The acetabulum is also subjected to degeneration, which positively correlates with age, but is undoubtedly the youngest age indicator, at least among the pelvic joints. It was introduced for the first time in 2004 (Rougé-Maillart et al., 2004). Since then, researchers have become interested in this anatomical part in relation to age estimation. Recently acetabulum has received increased attention (e.g. (Rissech et al., 2006, 2007, 2019; Rougé-Maillart et al., 2007, 2009; Calce, 2012; San-Millán et al., 2017)); however, the literature is not yet as comprehensive as for other skeletal indicators. Rissech and colleagues proposed a component-based approach using Bayesian inference, which is based on seven variables (chosen based on previous observations); each of them is scored separately (Rissech et al., 2006, 2007). To make the method better applicable for users, the computer program IDADE was developed. The reported difference between actual and estimated age was ± 10 years for almost 90% of individuals. A validation study (Calce and Rogers, 2011) pointed out that the accuracy of the Rissech method depends on the geographical distance

of the tested population from the reference population. The method of Rissech et al. was modified by Calce (2012), who proposed a simplified version with an extended application for females (Calce, 2012). Calce scores only three characteristics based on which individuals are assigned into one of the three broad age groups. However, the simplified approach had significantly worse accuracy than was declared by the original study (81%). Navega and colleagues, for instance, reported only 60% of correctly assigned individuals (Navega et al., 2018b) and Mays correctly assessed only 45% of examined individuals (Mays, 2014). This failure is attributed to the differences between the reference and test files (Mays, 2014; Navega et al., 2018b). Rougé-Maillart and colleagues combined the auricular surface (four traits) and the acetabulum (three traits), and developed a composite-based system (Rougé-Maillart et al., 2007, 2009).

Generally, the above-mentioned methods usually performed worse in other references samples (than in the ones used for developing the particular method).

In 2016, San-Millán and colleagues revised the Rissech method (2006) and redefined three of the seven variables (San-Millán et al., 2017). To date, this revised method has been applied and evaluated on another population sample only once, and by the same author's team (San-Millán et al., 2019). The authors confirmed the applicability of the revised method in different population samples and reached significantly higher accuracy than with the original method of Rissech (2006).

According to these studies, acetabulum is a promising indicator for estimating the age of adults, especially individuals over 60 years of age. However, researchers should be aware of the geographical distance between the reference and tested populations.

The most popular and best-known methods of age-at-death estimation based on pelvic joints are summarized in Table 4 based on their references in recent (published since 2010) forensic and anthropological manuals.

**Table 4.** An overview of age-at-death methods using joints of *os coxae* (pubic symphysis, auricular surface, acetabular surface) according to their representation in selected bioarchaeological and forensic manuals.

| | **Dirkmaat (2012)** | **İşcan and Steyn (2013a)** | **Christensen et al. (2014c)** | **Langley and Tersigni-Tarrant (2017)** | **Nikita (2017c)** | **Latham et al. (2018)** | **Adserias-Garriga (2019)** | **Notes** |
|---|---|---|---|---|---|---|---|---|
| **Pubic symphysis** | | | | | | | | |
| McKern and Stewart, 1957 | Yes | Yes | Mentioned | Yes | Yes | Yes | Mentioned | 3 components; male |
| Gilbert and McKern, 1973 | Yes | Yes | Mentioned | Yes | Yes | Yes | Mentioned | 3 components; female |
| Brooks and Suchey, 1990 | Yes | Yes | Yes | Yes | Yes | Yes | Yes | 6 phases |
| Hartnett, 2010 | Mentioned | No | Mentioned | Yes | Mentioned | Yes | Yes | 7 phases |
| Slice and Algee-Hewitt, 2015 | Not published | Not published | Not published | No | No | Yes | Yes | Computational approach |
| **Auricular surface** | | | | | | | | |
| Lovejoy et al., 1985 | Yes | Yes | Yes | Yes | Yes | Yes | Yes | 8 phases |
| Buckberry and Chamberlain, 2002 | Yes | Yes | Mentioned | Yes | Yes | Yes | Yes | 6 components |
| Osborne et al. 2004 | Yes | Yes | Mentioned | Yes | Yes | Yes | Mentioned | 6 phases |
| **Acetabulum** | | | | | | | | |
| Rougé-Maillart et al., 2004, 2006, 2009 | No | Yes | No | Yes | Yes | Yes | Yes | 7 components; with auricular surface |

**Table 4.** *Continued.*

| | Dirkmaat (2012) | İşcan and Steyn (2013a) | Christensen et al. (2014c) | Langley and Tersigni-Tarrant (2017) | Nikita (2017c) | Latham et al. (2018) | Adserias-Garriga (2019)[1] | Notes |
|---|---|---|---|---|---|---|---|---|
| **Acetabulum** | | | | | | | | |
| Rissech et al. 2006, 2007 | No | Yes | No | Yes | Yes | Yes | Yes | 7 components |
| Calce, 2012 | No | Yes | No | Yes | Yes | Yes | No | Evaluation of 3 variables into 3 age categories |
| San-Milán et al., 2016 | Not published | Not published | Not published | No | No | No | No | 7 components |

Mentioned – methods are not described in detail, only reference is given
Not published – the method has not yet been published at the time the manual was published

***Conclusion***

Traditional anthropological methods based on the evaluation of gross morphology seem to be less accurate and reliable in comparison to certain biochemical methods (Ritz-Timme et al., 2000; Adserias-Garriga and Zapico, 2018). They are, however, much more accessible and cheap for the vast majority of researchers, institutions and universities. Morphological methods generally underestimate older individuals and overestimate younger individuals. Especially the age estimation of the eldery is very inaccurate, ending with an the open-ended intervals (Milner et al., 2018).

However, articulation surfaces of *os coxae* show differences in aging patterns, offering their chronological combination as age indicators without neglecting their biological ability to reflect age-related changes. The pubic symphysis shows surface changes correlated to chronological age up to approximately 40 years; the acetabulum presents age-related changes in even older individuals (over 60 years), and the auricular surface of the ilium shows the changes approximately between the aforementioned indicators. All are located very close to each other, on one skeletal element, *os coxae*.

Nevertheless, the choice of skeletal indicators and methods is always determined by the nature of the remains to be examined, by their preservation and by the particular situation, e.g. their importance when it is a forensic case. The best possible approach is to assess the skeleton as a whole

## 3.2. Effects of external factors on skeletal senescence changes

One of the main reasons why there is no strict linear relationship between chronological and biological age is that bone degeneration used to be influenced by many different factors (internal and external) with a highly variable rate. These factors may affect age-at-death estimation since they manifest as non-age-related degenerative changes. Adult individuals are more affected by external factors than juveniles because they are exposed to them for a longer time (Christensen et al., 2014b). Unfortunately, the variability in age-progressive changes is found not only at the level of individuals within the same population, but also among different populations (Ferembach et al., 1980; Mays, 2015). Given that, two individuals of the same age may show completely different morphological changes. It is due to different activity levels, habits, lifestyle, etc. within and among populations (Langley et al., 2017).

The external factors have an effect on the processes of bony surface degeneration due to bone plasticity. Therefore, factors like physical activity, body mass, body size, pathological condition (e.g. osteoarthritis), occupation or obesity (e.g. (Cunha et al., 2009; Merritt, 2015, 2017a; Wescott and Drew, 2015; Calce et al., 2018a; b)) could have considerable impact on the senescence rate and represent a source of variability. Mays even states that most variation (60%) in skeletal indicators used for adult age estimation is caused by factors other than age (Mays, 2015).

### *Joint diseases*

A pathological condition is one of the factors that influences the degenerative changes and so may have an effect on the age estimation methods. Bone morphology could be affected by different pathologies. Some of them can be easily detected; some may cause only slight changes, potentially interchangeable with normal aging patterns (Rissech et al., 2004, 2018).

Recently, attention has been paid to the effect of joint disease osteoarthritis (OA) on age-at-death estimation, particularly in weight-bearing joints (Calce et al., 2018a; b; Winburn and Stock, 2019). Along with osteoporosis, they are often present in advanced-age individuals (Adserias-Garriga and Wilson-Taylor, 2019). OA is a progressive, degenerative disease which is accompanied at the beginning by the degradation of articular cartilage and later possibly followed by bone-on-bone contact, which can induce reactive bone formation (Loeser, 2010; Galasso et al., 2012; Mobasheri and Batt, 2016). Development of osteophytes is often considered as a form of degenerative bone change, thus as one of the indicators of the aging process. Osteophyte presence is just one of the features that makes OA recognizable (Rogers et al., 1990; Van der Kraan and van den Berg, 2007). In the case of the pelvic bone, it is used in several methods (Rissech et al., 2006; Rougé-Maillart et al., 2009; Calce and Rogers, 2011; Calce, 2012).

The fact that osteoarthritis is one of the most common joint diseases in today's world (Arden and Nevitt, 2006) with considerably underestimated (Jafarzadeh and Felson, 2018), yet growing prevalence (Arden and Nevitt, 2006; Jafarzadeh and Felson, 2018), it should not be overlooked. Further increase in its incidence can be expected in the future and there are valid concerns about whether OA can affect current age estimation methods based on the articulation surfaces of the pelvic bone (Calce et al., 2018a). The study of

Calce and colleagues revealed, firstly, that OA does affect the accuracy of age estimation for each of the three *os coxae* joint surfaces, with most noticeable error in young individuals exhibiting signs of osteoarthritis (Calce et al., 2018b). Secondly, joints of the *os coxae* demonstrate a variable pace of arthritic progression, but all of them seems to be consistently older.

The formation characteristic of OA is undoubtedly related to age, but its use for the aging of skeletal remains is very limited, since the etiology of OA itself is multifactorial and the degree of variation is too high (Aykroyd et al., 1999; Garvin et al., 2012; Alves-Cardoso and Assis, 2018). The use of OA occurrence is only recommended as supplementary information in an effort to estimate age-at-death (Aykroyd et al., 1999; Alves-Cardoso and Assis, 2018).

Another pathological condition that may have some effect on age estimation is DISH (Diffuse Idiopathic Skeletal Hyperostosis). However, no significant effect of DISH with the use of the Rissech acetabular method (Rissech et al., 2006) was revealed (Mays, 2012). However, these results must be viewed critically with respect to the very small sample of individuals with DISH, and further research is needed.

There are many other pathologies that could influence bone morphology; for example, rheumatoid arthritis, juvenile chronic arthritis, ankylosing spondylitis, psoriatic arthritis, Reiter's syndrome, and tuberculosis (Rissech et al., 2004). However, their influence on the morphology of age-related indicators of *os coxae* is also not very well explored. Only little influence of bone loss in the three skeletal indicators of *os coxae* was reported (Rissech et al., 2018).

***Obesity, body mass and body size***

Over the past few years, the increasing prevalence of obesity and its possible impact on age-related changes in the pelvic joints have been reflected in anthropological literature (e.g. (Merritt, 2015; Wescott and Drew, 2015; Winburn, 2018)). All three articular surfaces of *os coxae*, the pubic symphysis, the auricular surface and the acetabulum, participate in body mass transfer from the upper part of the body to the lower limbs. These weight-bearing joints are heavily loaded, and the joints of obese individuals are exposed to even greater biomechanical stress (Browning and Kram, 2007; Berenbaum and Sellam, 2008). Therefore, it is assumed that obese individuals, particularly younger

ones, will be overaged. Moreover, obese individuals exhibit biomechanical differences during locomotion (Ohtake, 2008; Sarkar et al., 2011). Thus, it is very probable that the rate of skeletal aging (joint degeneration) will be affected in obese individuals (in the sense that obese individuals will appear older). Indeed, these assumptions are consistent with the conclusions of a recent study of Wescott and Drew, where an over-age tendency was observed in obese individuals. The age estimation of iliac auricular surface is influenced by obesity more markedly, while the pubic symphysis seems more resistant (Wescott and Drew, 2015). The acetabulum was also found resistant to obesity effects (Winburn, 2018).

Similar concerns (but with opposite effect – slower bone remodeling rate) were raised about small and light-weight (with low body mass) individuals, and consequently proved. Small individuals and those with low body mass have decelerated skeletal aging and appear to age slowly (Merritt, 2015; Campanacho, 2016). It is therefore appropriate to ask, especially in the forensic context, how obesity or body mass affects joint morphology, as well as if it affects the accuracy and reliability of aging methods (Merritt, 2015, 2017b). Taking these factors into account, the adjusting of reported age range should be considered (Wescott and Drew, 2015). Nevertheless, research using documented osteological collections is limited as to its information about the weight of individuals. Such information is usually very rare, if available at all. They are, for example, from the time of death and may not reflect the condition of the individual during life.

### *Physical activity and occupation*

Apart from obesity and body mass, it could be intuitively assumed that physical activity will also affect the progression of age-related changes of *os coxae* joints. Therefore, more a physically demanding lifestyle, occupation or activity will correspond to a greater rate of joint degeneration and will be responsible for influencing skeletal age indicators. Nevertheless, only a few studies have addressed this issue (Campanacho et al., 2012; Mays, 2012; Miranker, 2017; Winburn, 2018), but none of the studies on the pelvic bone joint indicators have confirmed the relationship between age and activity or occupation.

The number of difficulties with occupation must be mentioned. The first issue concerns the record of the occupation being performed during life. Of the identified osteological collections, we usually only have information of the last occupation, but we

do not know how long it had been done (Mays, 2012). Moreover, occupation does not represent all the physical activities during an individual's life (Mays, 2012, 2015), so it is very difficult to distinguish among factors such as diet, nutrition, lifestyle and other external factors (Mays, 2015).

## 3.3. Limitations to skeletal aging and new challenges

***Population specificity of aging methods***

Various groups of people across the world (individuals living in different geographical areas) are affected by internal and external factors to varying degrees. That may affect the rate at which the body naturally goes through a process known as 'wear and tear' and thus it may also affect the manifestation of age-related changes observed on different parts of the skeleton (e.g. (Hoppa, 2000; Olze et al., 2004; Ubelaker, 2018b; Ubelaker and Khosrowshahi, 2019)). Therefore, it is not surprising that several studies on various population samples have confirmed that the methods they applied in their sample did not perform as well as stated by the original studies in their reference sample (e.g. (Schmitt et al., 2002; Schmitt, 2004; Bassed et al., 2011; Buk et al., 2012; Mays, 2014; Jooste et al., 2016; Navega et al., 2018b)). Information about the population variability of aging rates are, however, inconsistent, since no differences were revealed in some studies (Sakaue, 2006; Kimmerle et al., 2008).

While we often hear about population-specific aging rates and the need to develop population-specific aging methods, there is also great intra-population specificity. The question is whether it makes sense to continue in developing population-specific (or group-specific) methods when there is so much variability within populations (groups) themselves. This is maybe the case when it would make sense to develop methods based on the analysis of age-related changes in restricted ecogeographical areas (Kim, 2016), or in a large multi-population (group) sample, as was done by, for example, Schmitt (2002), which would cover much more variability of aging processes. The former could be explained as a group of people (aggregates) who cover a range of variability that is influenced by the specific region or geographical location they share (Franklin and Flavel, 2019). Such an approach was used for age-at-death estimation proposed by Kim in her dissertation thesis on a pooled Asian sample which outperformed specific models (Kim, 2016). On a similar basis with the use of a pooled sample from a specific geographical

area, an approach to estimate stature using the tibia was developed. A multi-population sample consisting of populations from four continents was used, for example, by Schmitt (Schmitt, 2001; Schmitt et al., 2002). In both aforementioned cases, data were evaluated visually. Approaches based on more populations or groups from a specific area are more robust than single population-based approaches. In this era of globalization (Spradley et al., 2008; L´Abbé and Steyn, 2012) when uncertainty about the origin of skeletal remains increases significantly, these approaches would be particularly appropriate.

The idea of proposing aging methods based on multi-population data or data pooled from specific geographical regions is quite new and more research is required.

***Methodological limitation – single or multiple skeletal indicators***

For a very long time, there was a fixed idea that a single indicator could capture the entire adult period of human life. To this day, we have plenty of single-indicator methods that try to capture age-related changes of the particular skeletal indicator from the point of an individual reaching maturity to the end of his life (e.g. (Brooks and Suchey, 1990; Buckberry and Chamberlain, 2002; Rissech et al., 2006)). However, accurate and reliable age estimation with single-indicator methods is not possible. It is very slowly being accepted, and the original idea abandoned, that various indicators are more suitable and contribute to age estimation the most in different periods of human life (e.g. (Martrille et al., 2007; Milner and Boldsen, 2012a; Adserias-Garriga and Wilson-Taylor, 2019)). This fact was emphasized in *Chapter 3.1.* for the pelvic bone articulation areas.

Interestingly, the idea of multifactorial or complex age estimation is not entirely new. Nonetheless, the Complex method of Acsádi and Nemeskéri suppose equal informative potential of all involved indicators (Acsádi and Nemeskéri, 1970). Some authors argue that using more varied anatomical structures (multifactorial approach) at the same time, combined and statistically evaluated into one estimate will bring more accurate results (e.g. (Lovejoy et al., 1985a; Bedford et al., 1993; Baccino et al., 1999; Anderson et al., 2010)), because different indicators can provide a more balanced age estimate. Others prefer using many different aging methods whose results they combine into wide ranges (Falys and Lewis, 2011; Garvin and Passalacqua, 2012). For a long time, there has not been consensus on how to combine multiple indicators to obtain a single estimate with reasonable (not too wide) confidence intervals, as well as which methods should be

employed (Franklin, 2010; Garvin et al., 2012; Milner and Boldsen, 2012b). The approaches using as many skeletal and dental indicators as possible to achieve the highest accuracy are currently recommended by some researchers (e.g. (Cunha et al., 2009; Ubelaker, 2018b; Adserias-Garriga and Wilson-Taylor, 2019)). However, even though these methods could be statistically robust, they often perform comparably with single-indicator methods (e.g. (Saunders et al., 1992; Schmitt et al., 2002; Bethard, 2005; Martrille et al., 2007)).

***Methodological limitations – statistics***

Statistical calculations used to estimate age-at-death can be a great source of bias, so they are often criticized. If we skip the very first methods (Todd, 1920; Lovejoy et al., 1985b), most aging procedures are based on linear regression that is used to correlate chronological age and the morphological score of the indicator (e.g. (Katz and Suchey, 1986; Buckberry and Chamberlain, 2002; Falys and Prangle, 2015)). In the regression-based models, an equation of regression line is used to get the predicted ages (Aykroyd et al., 1999). The main problem is that they assume a linear relationship between morphological traits and the chronological age which is not confirmed (Schmitt et al., 2002). These approaches tend to overestimate the age of the young and underestimate that of the old (Aykroyd et al., 1999), a problem also known as 'attraction to the middle' (Masset, 1990). The biased estimates may be one of the reasons why we have a low number of older individuals in archaeological populations (Konigsberg and Frankenberg, 1992). We simply do not yet have methods to detect them properly.

The use of linear regression is also burdened with a phenomenon called 'age mimicry', first highlighted by Bocquet-Appel and Masset (Bocquet-Appel and Masset, 1982), where age estimate distribution of the test sample was contaminated (biased) by the age distribution of the reference sample on which the method was developed (Konigsberg and Frankenberg, 1992; Boldsen et al., 2002).

The current situation with adult aging techniques has resulted in broad overlapping or open-ended age intervals, for which they are often criticized (e.g. (Brooks and Suchey, 1990; Boldsen et al., 2002; Buckberry and Chamberlain, 2002)). Several researchers have stated that accurate estimation is possible only into three wide age intervals (e.g. (Schmitt et al., 2002; Buk et al., 2012; Calce, 2012; Adserias-Garriga and Wilson-Taylor, 2019));

for example, less than 30, 30–60, 60 years and more (Buk et al., 2012); 20–34 years, 35–45 years, and 46+ years (Falys and Lewis, 2011); or, for instance, up to 40 years, 40–65 years, and over 65 years (Adserias-Garriga and Wilson-Taylor, 2019). It is better to place an individual within a wide age range than try to fit him incorrectly into narrower intervals. On the other hand, wide intervals are not very helpful in narrowing the range of possible matches, especially in medico-legal identification. However, even the three broad age ranges do not guarantee more accurate age estimates, as shown by some studies validating the method of Calce from 2012 (Mays, 2014; Navega et al., 2018b). The original efforts to estimate the age of adults within ten-year or even five-year age intervals (Lovejoy et al., 1985b) often led to very inaccurate age estimates; for example, only 33% of correctly aged individuals (Osborne et al., 2004). This was also recently confirmed by Buk and co-workers (2012). Soon it was found that it is not realizable because the variability of the aging process does not allow it and the error is too high.

As the regression-based models have been shown to be inappropriate for such complex processes as skeletal aging (Lucy et al., 1996; Schmitt et al., 2002; Miranker, 2016), the application of probabilistic approaches seems much more appropriate and promising. The Bayesian approach, in contrast to regression-based approaches, reduces the systematic error as well as the trend of under- and over-aging; it is less sensitive to the effect of age mimicry and provides better age estimates (closer to the actual age) and smaller confidence intervals. Given that fact, this approach of converting age indicators into age estimates has received much attention (Lucy et al., 1996; Aykroyd et al., 1999; Boldsen et al., 2002; Schmitt et al., 2002; Digangi et al., 2009; Buk et al., 2012; Brennaman et al., 2017; Kotěrová et al., 2018a), especially in response to the Rostock manifesto (Hoppa and Vaupel, 2002). The manifesto is the theoretical framework resulting from the paleodemographic workshop in 1999, which aimed mainly to integrate new biostatistical approaches, i.e. the Bayesian approach, into paleodemographic analyses, and particularly to adult age-at-death estimation (Hoppa and Vaupel, 2002). These approaches also soon became reflected in forensic age-at-death estimation (e.g. (Schmitt et al., 2002; Rissech et al., 2006; Steadman et al., 2006; Konigsberg et al., 2009)).

A possible solution of both the aforementioned points (statistical approaches and number of employed indicators) could be Boldsen's multifactorial technique – Transition analysis (TA), whose development began in 1996 and was published a few years later (Boldsen et al., 2002; Milner and Boldsen, 2012b). Boldsen's technique addresses several

problems accompanying age estimation and is considered as statistically appropriate approach (Jooste et al., 2016) that combines multiple indicators systematically. The approach combines three skeletal indicators (pubic symphysis, auricular surface and cranial sutures) that are scored with a component-based system. The Bayesian statistics along with the transition analysis, which computes the age of transition from one stage to another, is applied and all incorporated in the user-friendly ADBOU program. As an output, the researcher obtains a multifactorial likelihood estimate and a maximum likelihood estimate for individual skeletal structures with confidence intervals (Boldsen et al., 2002). The confidence intervals are 'individualized', which is preferred for single cases, because every skeletal remains shows a different degree of error (Boldsen et al., 2002). Standardized age categories may be better in bioarchaeology since they enable better between-group comparison.

A great advantage is that the TA method is applicable to fragmentary remains, thus, the scoring of all the features is not necessary, which is impossible with most phase and component methods. This approach could be applied to single cases as well as to archaeological samples. It uses archaeological and forensic prior probability functions. Furthermore, the TA should overcome age-mimicry, eliminate open-ended intervals and reduce the bias caused by different demographic profiles in the target and reference samples through the provision of informative prior. Nevertheless, the results of studies testing the validity of TA are inconsistent and ambiguous. Some studies affirmed that TA outperformed traditional single-indicator methods (Godde and Hens, 2015); whereas some did not confirm these conclusions and did not find the results satisfactory (Milner and Boldsen, 2012b; Xanthopoulou et al., 2018). Thus the problem with under- or over-estimation has not been resolved (Fojas et al., 2018; Hagelthorn et al., 2019). The TA method also performed differently among populations and sexes (Hagelthorn et al., 2019). Especially in the case of individual age estimates, the method performed worse than when the aim was to capture a general age-at-death distribution of a population (Milner and Boldsen, 2012b). The developers of the method admitted that TA did not outperform experience-based assessment (Milner and Boldsen, 2012b). On the other hand, Fojas and colleagues declared that the majority of actual ages fell within the confidence intervals (Fojas et al., 2018). Moreover, TA better captures older individuals of the population, who are mostly overlooked by traditional methods (Bullock et al., 2013). TA is used among researchers, though not widely (Falys and Lewis, 2011; Garvin and Passalacqua, 2012).

One of the reasons why it does not perform better could be the limitation of the skeletal indicators used in the analysis (Milner and Boldsen, 2012b). A new version of ADBOU, utilizing many more skeletal indicators is being prepared (pers. comm. Boldsen and Milner at workshop AAPA, Cleveland, 2019).

Very recently, Nikita and Nikitas (2019) proposed two techniques for regression analysis to take into account the demographic profile. They found that under certain conditions, these techniques outperform the Bayesian approach (Nikita and Nikitas, 2019). Apart from this approach, more attention has been paid to the statistics that are being applied to age estimation in the last few years. The focus of researchers was shifted to sophisticated data mining methods, namely, artificial neural networks, decision trees, nearest neighbours, the Sugeno Fuzzy integral and other computational intelligence methods (e.g. (Anderson et al., 2010; Buk et al., 2012; Martins et al., 2012; Navega et al., 2018a)). These advanced statistical approaches also penetrate other aspects of the biological profile, such as sex estimation (Santos et al., 2014; Navega et al., 2015b; Langley et al., 2018), and ancestry assessment (Hefner et al., 2014; Navega et al., 2015a; Scott et al., 2016).

### *Conclusion*

As the previous text suggests, many small steps towards increasing the accuracy and reliability of age estimation methods have already been taken, but many are still waiting to be discovered.

The use of single-indicator methods does not allow accurate and reliable estimation of age, as one indicator alone does not reflect the entire adult period of life. At the same time, the process of combining multiple indicators is not sufficiently standardized and these methods often result in even wider age intervals.

The rate of degenerative changes can be influenced by various external, as well as internal factors, and ecogeographical origin. Therefore, it can vary among different groups of people, which should be taken into account when applying age estimation methods that are created from distant reference samples. In the future, attention should be paid to a better understanding the issue of population or group-specific rates of the aging process.

Finally, very often it is the case that the methods show accurate results in the reference population (the population in which the methods were developed) but fail when they are applied to different populations, thus resulting in low reliability. Currently, it

seems that no matter which approaches are performed, all efforts for accurate estimates mostly end in wide age intervals, usually within three broad intervals; for example, less than 30, 30–60, and 60+ years. The use of Boldsen's transition analysis (in its current form) is not yet well established.

# 4. AGE-AT-DEATH ESTIMATION: NEW APPROACHES AND INSIGHTS

Given that there are increasingly higher standards of evidence required to be met in medico-legal cases, i.e. Daubert's requirements, researchers are pushed to produce more objective results. Thus, they have turned to imaging and scanning techniques, which have become more accessible through rapid technological advances. These promise, among many other things, the objectification of the evaluation of senescence changes for age-at-death estimation. The first section (*Chapter 4.1.*) is briefly focused on the advantages of virtual anthropology and shows the differences among the 3D models obtained through surface scanners and computed tomography. Finally, it presents examples of the use of virtual anthropology in the estimation of biological profiles. As age-at-death estimation is the main objective of this thesis, the next part is specifically devoted to the age-related changes taking place on the *os coxae* assessed in the virtual environment (*Chapter 4.2.*). The last section (*Chapter 4.3.*), which concludes the theoretical part, focuses on the quantitative analysis of age changes in the joint surface.

## 4.1. Imaging technologies: The rise of virtual anthropology

Medical imaging technologies (computed tomography, ultrasonography, magnetic resonance imaging, radiography) and surface scanning technologies give anthropology a whole new dimension and in many ways offer advantages that are of great benefit. First of all, there is the possibility of creating and archiving digital copies of human remains that can be preserved even if the actual template no longer exists. These 3D copies, whether with internal structure (from computed tomography) or surface models only, can then be easily transferred and shared among research institutions, which makes it possible to create large databases. In addition, digital skeletal collections allow for manipulation in a virtual environment, reducing the need to work with physical remains. Recently, 3D printing technologies for exhibiting or teaching purposes have been on the rise (e.g. (Friess, 2012; Klein et al., 2014; Cantín et al., 2015; Viggiano et al., 2015; Errickson, 2017; Guydish and Henson, 2017; Seguchi et al., 2019)). Finally, completely new analyses can be applied to the 3D records of human remains, enabling better monitoring of shape and morphological changes (e.g. (Villa et al., 2015a; Musilová et al., 2019)). These analytical approaches

(e.g geometric morphometrics) open up a new perspective on biological profile estimation that cannot be provided by traditional visual or metric approaches performed on dry bones. One of the main advantages is the possible limitation or complete reduction of subjectivity that now burden conventional methods.

With the advent and adaptation of imaging methods for anthropology, researchers primarily employed computed tomography (CT) to visualize the desired skeletal structure in order to estimate the biological profile. 3D reconstructions based on CT are beneficial not only because they can be collected from recent populations (Ramsthaler et al., 2010; Colman et al., 2019), but also because they give non-invasive insight into the internal structures. Later, studies using surface scanners, whether based on laser or structured light technology, began to be integrated. They are much more affordable compared to CT, and do not involve any radiation or the need for a highly professional operator; in many cases they are easily portable and, last but not least, they also enable the recording of texture (e.g. (Friess, 2012; Kuzminsky and Gardiner, 2012; Seguchi and Dudzik, 2019)). However, it might be time-consuming to create 3D surface images for beginners and it requires some training (Seguchi et al., 2019). The surface scanner in many cases can provide more detailed information of surface, especially when compared to the CT-based model acquired from living people, where the resolution is lower due to the radiation dose, which must be kept to a minimum.

The virtual representations of human skeletal remains since then have proved to be useful and able to replace dry bones in biological profile analyses. They are, for example, commonly used for morphological or metric assessment of the individual's sex (Dedouit et al., 2007; Grabherr et al., 2009; Ramsthaler et al., 2010; Decker et al., 2011; Chapman et al., 2014; Mesteková et al., 2015), with generally positive results supporting the integration of virtual models into anthropological practice. However, caution is needed since not all sexing methods can be applied to virtual models (Colman et al., 2019). The consistency between the dimensions measured from dry and virtual bones has been established several times (Citardi et al., 2001; Verhoff et al., 2008; Corron et al., 2017), although some authors admitted small differences (Dedouit et al., 2007; Mullins and Albanese, 2017; Kotěrová et al., 2019). However, the important point is whether these differences have a significant effect on further analyses as biological profile estimation. Apart from the sex assessment analyses in the virtual environment, analyses of ancestry assessment or population

differences (e.g. (Cavaignac et al., 2017; Murphy and Garvin, 2018; Musilová et al., 2019)) are also performed.

## 4.2. The use of 3D models for age-at-death estimation

Virtual replicas of different skeletal anatomical structures reconstructed from a sequence of computed tomography images or acquired by 3D surface scanners could be used for age estimation in different ways. The extent of their use ranges from the basic applications (such as conventional visual methods derived from dry bones to their virtual replicas), through metric variables and characteristics of internal structures, to more sophisticated approaches based on mathematical quantification of joint surface changes. The following overview is specialized in 3D models of *os coxae*.

***Combining traditional methods with 3D data***

The very first and simplest experiments on estimating age applied traditional visual methods, originally developed on dry bones, to 3D models of the pelvic bone (e.g. (Telmon et al., 2005; Barrier et al., 2009; Biwasaka et al., 2013; Villa et al., 2013a; Merritt, 2018a; b; Hisham et al., 2019; Pattamapaspong et al., 2019)). Many studies assessed the reliability of CT reconstructed models for age estimation and recognized that the methods they tested, could not be fully applied to 3D CT models, because some features were scored with difficulty or improperly, and some could not be scored at all. Such features most often include: (micro)porosity, bone weight, surface texture and transverse organization (Barrier et al., 2009; Villa et al., 2013a; Merritt, 2018b; a; Pattamapaspong et al., 2019). Among the publications using CT technology, only few researchers reported excellent results of method performance (e.g. (Telmon et al., 2005; Hisham et al., 2019)). Some authors suggest modifying the existing macroscopic methods to make them more transferable to 3D reconstructions (Merritt, 2018a; b).

Villa and colleagues and Biwasaka and colleagues were the first (to my knowledge) who applied existing aging methods on 3D models acquired by laser scanner (Biwasaka et al., 2013; Villa et al., 2013a). Both of them obtained significantly better resolution than from the CT scanner, which made the morphological features easier to see. On the other hand porosity and pitting were still difficult to distinguish, if at all, and therefore could not

be evaluated. Nevertheless, even with laser scanners, the results of applying visual methods to 3D models are inconsistent, or depend on the method. Biwasaka reported no discrepancies in assigning the pubic symphysis phase (Biwasaka et al., 2013), which is in concordance with the results of Villa for the same articular area and method used (Villa et al., 2013a). On the contrary, this is not true for the auricular area and the Buckberry and Chamberlain method (2002), for which modifications are recommended if it is to be applied to virtual material (Villa et al., 2013a).

In conclusion, the traditional methods were developed on dry bones through their direct examination and thus work the best with them. However, virtual models can provide a useful substitute for dry bones when they, for instance, are not available. It must also be noted that even in the virtual environment, traditional methods are still burdened with subjectivity.

***Linear measurement and internal bone structure***

Given that CT data contain information about internal structures, they can be used and studied in relation to the aging process when bone destruction is not needed. The virtual environment enables the measuring of different structures and variables (i.e. morphometric analyses), for example, on CT scan sections (e.g. (Dedouit et al., 2007; Ferrant et al., 2009; Chiba et al., 2014)). Some authors directly assess the internal structure of bone (Barrier et al., 2009; Grabherr et al., 2009; Villa et al., 2013b; López-Alcaraz et al., 2015). The internal structure can also be assessed through densitometry (DXA = dual-energy x-ray absorptiometry). Aging techniques are based on the fact that bone mineral density (BMD) naturally decreases with age, resulting in a correlating relationship (e.g. (Navega et al., 2018a; Botha et al., 2019; Dubourg et al., 2019)). These approaches could partially remove the subjectivity mentioned in the previous point. Generally, the results indicate that these approaches could be useful for age estimation; however, further research is required.

***Surface quantification***

Lastly, some attempts to estimate age accurately and reliably have shifted towards computational methods using surface changes analyses that are enabled with the use of geometric morphometry and other advanced mathematical approaches (e.g. (Biwasaka et al., 2013, 2019; Villa et al., 2015a; Stoyanova et al., 2017)). Very recently, the homologous modeling method with use of Principal component analysis has been performed on CT data of the whole *os coxae* to explain age-dependent morphological changes (Biwasaka et al., 2019). However, accurate results were not obtained when applied to an unknown sample. Nevertheless, other approaches focused mainly on surface curvature analyses of the given joint area and the surrounding area (Biwasaka et al., 2013; Slice and Algee-Hewitt, 2015; Stoyanova et al., 2015, 2017; Villa et al., 2015a). For that purpose, they used affordable surface scanners and reached promising results. Therefore, the use of surface scanners for the quantitative analysis of the articular surfaces of *os coxae* is the subject of the next chapter.

## 4.3. Surface scanners: A tool for quantitative analyses of surface changes

After engaging medical imaging techniques in anthropology, surface scanners soon became popular, offering many advantages over, for example, CT (mentioned above). The first attempts to quantify the morphological changes of joint surface in order to estimate age focused on the articulation surfaces of *os coxae*, mainly on the pubic symphysis and then on the auricular surface of ilium. They were based on curvature analysis performed on 3D models acquired by laser scanners (Biwasaka et al., 2013; Villa et al., 2015a). In the study of Biwasaka and colleagues, mean curvatures of every 5 $mm^2$ were calculated and, based on them, concave and convex maps were created. The results showed that curvature analysis distinguished the first two phases of the Suchey-Brooks method (1990): the mean surface curvature correlated with the actual age in both sexes, but later phases could not be discerned.

The curvature analysis of Villa, which included mean curvature, the lowest and highest 10% of the curvature values and the percent of convex and flat surface, was extended to the auricular surface of *os coxae* (Villa et al., 2015a). The algorithms were first applied to bone casts (for the Suchey-Brooks and the Buckberry and Chamberlain methods), where they showed higher correlations with age than when applied and tested on

pelvic bones from skeletal collections. Although these two pilot studies have not produced ground-breaking results and have failed to increase the accuracy and reliability of the age estimation, they are a very good basis for future research towards the objectification of age-at-death estimation.

At present, the most recent attempt to quantify age changes of the pelvic joint surfaces is the approach of Slice and Algee-Hewitt and Stoyanova and colleagues (Slice and Algee-Hewitt, 2015; Stoyanova et al., 2015, 2017). Together, these teams proposed a fully quantitative method for age estimation from the pubic symphyseal surface of male individuals from white North American populations. The proposed shape analysis method is based on two regression models that capture surface flatness. They are TPS/BE (Thin plate splines/Bending energy) proposed in Stoyanova et al. (2015) and SAH-score proposed by Slice and Algee-Hewitt (2015). One model (VC) measures the shape of the ventral margin of the pubic symphyseal surface (Stoyanova et al., 2017) and two multivariate models combine the previous ones (SAH+VC and TPS/BE+VC). The results were expressed as RMSE values (ranging between 13.68 and 16.55 years), bias (−1.82 and −2.73 years) and inaccuracy (10.79 and 12.86 years). These results are more or less comparable to traditional visual methods (e.g. (Brooks and Suchey, 1990)), suggesting this might be the right direction for research. All these models are incorporated into the free user-friendly ‘forAGE’ software, allowing other practitioners to test the new approach and compare their results.

### *Conclusion*

The penetration of new technologies (e.g. medical imaging techniques and surface scanning technologies) into anthropology is undoubtedly of enormous importance, not only for methods of estimating biological profiles. In the case of age-at-death estimation, they help with the objectification of data collection and evaluation of senescence manifestation. For example, we encounter the possibility of monitoring age-related changes in internal structures without destroying the particular skeletal structure. Quantitative analyses of the articulation surface have become even more attractive, but these approaches are still in their infancy. Nevertheless, they are showing very encouraging results and we can hope that further research will improve them.

The increasing interest in scanning technologies as they are implemented into daily anthropological practice (e.g. age-at-death estimation) raises questions about the comparability of outputs from various devices and their impact on subsequent biological profile analyses. Further research should exploit the potential of quantitative methods to assess surface changes in articular areas objectively by using sophisticated statistical approaches (as, for instance, the Bayesian approach, neural networks, decision trees, and random forests) to overcome the three aforementioned broad age intervals. Beyond that, research should focus on the involvement of various age indicators with regard to what information about age they can provide.

# Part II: Personal contribution to age-at-death estimation methods

## 5. THE AIMS OF THE RESEARCH

The aims of this dissertation are based on the previous text and the state of the issue of adult age-at-death assessment.

The research had several main objectives:

1) **Improving the evaluation of visually acquired data by applying sophisticated mathematical approaches**

The first aim is to apply various mathematical approaches to visually scored data in order to reach more accurate age estimation. We aim to divide the adult human lifespan into more than the three wide age ranges (less than 30 years, 30–60 and over 60) proposed by Buk et al. (2012), which have been used so far to estimate age accurately and reliably at the same time.

2) **Evaluating the impact of various surface scanners on quantitative age-at-death assessment**

With the introduction of scanning technologies into anthropology and their incorporation into biological profile analyses, such as age estimation, it is necessary to know the limits of these technologies. While 3D scanners are increasingly used in workplaces equipped with various devices, their impact on further analyses has not yet been sufficiently explored. We aimed to investigate the comparability of 3D models acquired through different scanning technologies – laser and structured light. Moreover, we tested the influence of possible differences in captured surfaces on the estimation of age.

### 3) Validating the aging method of Stoyanova et al. on non-American populations

The method of Stoyanova and colleagues (2015, 2017) is the only approach that offers user-friendly software for the fully quantitative assessment of age-related surface changes. The aim was to validate the approach in different populations (the original method was derived from an American sample), to assess its performance and possible shortcomings, and to find solutions in the form of a proposal for our own approach. A prerequisite for such a study is the creation of a database of 3D models of pelvic bones based on a metapopulation sample of individuals of known age and sex. In such a dataset, the inter-population variability of senescence processes is considered, which may lead to the mitigation of the effect of population specificity.

# 6. MATERIALS AND METHODS

## 6.1. Materials

### 6.1.1. A multi-population dataset of visually scored changes on the *os coxae*

The dataset consists of visually-evaluated age-related changes to two traditionally used skeletal indicators, the pubic symphysis and the auricular surface of ilium. They were assessed according to the scoring system developed by Schmitt (Schmitt, 2001, 2005; Schmitt et al., 2002); a detailed description is shown in Table 5.

**Table 5.** Summary table of the scoring system for the pubic symphysis and the sacro-pelvic surface of the ilium (Kotěrová et al., 2018a).

| **Variables** | **Scoring** | | | |
|---|---|---|---|---|
| | **Score 1** | **Score 2** | **Score 3** | **Score 4** |
| **Pubic symphysis (PUS)** | | | | |
| Posterior plate (PUSA) | Ridges and Furrows | Surface flattening, appearance of dorsal margin | Complete | - |
| Ventral plate (PUSB) | Ridges and Furrows | Surface flattening, rampart expanding | Ventral rampart completely formed | - |
| Dorsal lip (PUSC) | Regular dorsal rim | Plain lipping | - | - |
| **Surface sacro-pelvic (SSPI)** | | | | |
| Transverse organization (SSPIA) | Undulation or striae present | Undulation or striae absent | - | - |
| Texture and porosity (SSPIB) | Dense surface | Uniform or partially coarsening granular texture | Coarse granular texture, porosities | Irregular surface, deep porosities, bone destruction |
| Apical activity (SSPIC) | Apex sharp, margins regular | Apex broad, rim formation | - | - |
| Retroauricular activity (SSPID) | Smooth surface | Irregular surface, osteophytes | - | - |

The skeletal remains that were studied were the *ossa coxae* (n=941) of adult individuals of both sexes between 17 and 100 years. The remains originated from nine osteological collections, namely the Coimbra Identified Skeletal Collection CISC (Cunha and Wasterlain, 2007) from Portugal; the Simon Identified Skeletal Collection from Switzerland (Perreard Lopreno, 2007; Abegg and Desideri, 2018); the Spanish Collection (Schmitt, 2001); Euro-American and Afro-American samples from the Hamman-Todd Collection (Mensforth and Latimer, 1989); African samples (Afrikaner, Zulu, and Soto) from the Dart Collection (Dayal et al., 2009); and, lastly, the Chiang Mai Collection (Traithepchanapai et al., 2016) from Thailand. The detailed composition of the osteological collections, including the origin of each sample and the number of male and female individuals, is given in Table 6.

The whole dataset was collected and kindly provided by Aurore Schmitt.

**Table 6.** The composition of a multi-population dataset (Kotěrová et al., 2018a).

| Sample | Ancestry | Origin | Collection | Females | Males | Total |
|---|---|---|---|---|---|---|
| Portugal | Caucasian | Europe | Coimbra | 85 | 74 | 159 |
| Switzerland | Caucasian | Europe | Geneva | 17 | 26 | 43 |
| Spain | Caucasian | Europe | Madrid | 34 | 33 | 67 |
| USA 1 | Caucasian | America | Cleveland | 73 | 69 | 142 |
| USA 2 | Afroamerican | America | Cleveland | 40 | 41 | 81 |
| Afrikaner* | Caucasian | Africa | Johannesburg | 32 | 35 | 67 |
| Zulu* | African | Africa | Johannesburg | 102 | 102 | 204 |
| Soto* | African | Africa | Johannesburg | 44 | 33 | 77 |
| Thailand | Asian | Asia | Chang-Mai | 45 | 56 | 101 |
| | | | Total: | 472 | 469 | 941 |

*South Africa

### 6.1.2. Czech medieval sample of 3D models of the *ossa coxae*

Part of our research was also to compare the outputs from various scanning devices and find out if any differences in surface capture affect the estimation of the biological profile. As all the necessary scanners could not be transported to the osteological collections described above, we selected a small sample of pelvic bones of the medieval

population that could be scanned by the selected devices in two cities in the Czech Republic (Prague, Pilsen).

The skeletal sample represents a medieval population from the cemetery of the second church in the Mikulčice settlement ($9^{th}$–$10^{th}$ century AD) in South Moravia (Poláček, 2008). It consists of the well-preserved *ossa coxae* of adult individuals (n=29), some of which are represented by both pelvic bones. The age and sex of the individuals are not known, but this is not necessary for our purposes, since the aim was to compare the resulting estimates and not to verify the actual age and sex.

### 6.1.3. Multi-population dataset of the 3D digital joint surfaces of the ossa coxae

During the years 2016–2019, a large dataset of 1268 surface models of adult *ossa coxae* was created. This dataset consists of four European identified skeletal collections and one Asian collection. All of them are described below. The skeletal remains of adult (17–101 years) males and females were selected for digitization. The *ossa coxae* with obvious pathological changes were excluded. The left-sided *ossa coxae* were preferred; however, the right-sided were included as well in order to be able to evaluate the asymmetry in the aging process and to increase the biological variability. When both *ossa coxae* of the same individual were available for the particular analysis, they were treated independently. Table 7 shows the composition of each osteological collection, including the numbers of the left and right-sided *ossa coxae* (n1) and the numbers of individuals from which at least one *os coxae* was digitized (n2), mean age and median of age for each sex separately and together.

#### *The Coimbra Identified Skeletal Collection (CISC)*

The CISC (hereinafter Coimbra 1) is stored at the University of Coimbra, Portugal, and consists of 505 individuals of which 45 are juveniles. The collection was put together by Professor Tamagnini and is made up of skeletons from the main Coimbra cemetery. The collection consists of individuals who lived in the $19^{th}$ and $20^{th}$ centuries; they were born between 1817–1924 and died between 1904 and 1938 (Cunha and Wasterlain, 2007).

A total of 285 *ossa coxae* from CISC were digitized in 2016, representing 148 individuals.

### *The 21st Century Identified Skeletal Collection (CEI/XXI)*

This is a modern still-growing collection (hereinafter Coimbra 2) comprising of unclaimed skeletons from the cemetery, also stored at the University of Coimbra. This collection is represented by individuals who lived mostly in the 20th and 21st centuries. At the time of studying the collection (autumn 2016), all the individuals had died between 1995 and 2009. At that time, the collection contained around 250 individuals (Ferreira et al., 2014).

A total of 235 *ossa coxae* of 132 individuals were scanned in 2016 as well.

### *The Heraklion collection*

The Heraklion collection is housed at the facilities of the Forensic Pathology Division of the Hellenic Ministry of Justice and Human Rights in Crete, Greece. This collection consists of skeletal material from Cretan cemeteries in Heraklion and is represented by around 200 individuals. The representatives lived in the 19th and 20th centuries; they were born between 1867 and 1956 and died between 1968 and 1998 (Kranioti et al., 2008; Kranioti and Michalodimitrakis, 2009).

A total of 202 *ossa coxae* were recorded from the Heraklion collection representing 99 individuals. Unfortunately, the actual age-at-death for sixteen individuals is not known from this sample, so they could not be used for age assessment in age-at-death estimation studies. Future use of these skeletal remains is not excluded, for example, for sex estimation. Eventually, a total of 170 bones (88 individuals) could be used to estimate the age-at-death. The collection was digitized in 2017.

### *The Simon Identified Skeletal Collection*

The Simon collection was gathered from cemeteries of the canton of Vaud and now is housed at the Laboratory of Prehistoric Archaeology and Anthropology of the University of Geneva, Switzerland. It comprises 495 individuals in total, who died between the late 19$^{th}$ century and the first half of the 20$^{th}$ century (Perreard Lopreno, 2007; Abegg and Desideri, 2018).

In total 277 *ossa coxae* representing 161 individuals were scanned. The Simon collection was digitized in 2018.

### *The Khon Kaen University Collection (KKU)*

The KKU collection is stored at the Department of Anatomy, Faculty of Medicine of Khon Kaen University. The collection is comprised of donated bodies through the medical school's body donation program, and is still growing. The program began in 1973 and donations were made mostly between 1988 and 2016 (Techataweewan et al., 2017b; a). At the time of studying the collection (at the beginning of 2019), it numbered over 1000 individuals, who had lived between 1908–1988 and died between 1988–2015.

A total of 301 *ossa coxae* (235 individuals) were digitized from KKU in 2019.

**Table 7.** A summary table of the composition of the osteological collections.

| | Sex | n $_1$ | n $_2$ | Mean age $_1$ | Median $_1$ |
|---|---|---|---|---|---|
| **Coimbra 1** | **Male** | 148 | 77 | 45.03 | 45 |
| | **Female** | 137 | 71 | 49.82 | 50 |
| | **Total** | 285 | 148 | 47.33 | 47 |
| **Coimbra 2** | **Male** | 104 | 60 | 65.60 | 69.5 |
| | **Female** | 131 | 72 | 79.81 | 83 |
| | **Total** | 235 | 132 | 73.52 | 78 |
| **Heraklion** | **Male** | 83 | 43 | 69.05 | 71 |
| | **Female** | 87 | 45 | 72.23 | 74 |
| | **Total** | 170 | 88 | 70.68 | 72 |
| **Geneva** | **Male** | 142 | 79 | 47.37 | 46 |
| | **Female** | 135 | 82 | 49.16 | 48 |
| | **Total** | 277 | 161 | 48.25 | 47 |
| **Khon Kaen** | **Male** | 175 | 136 | 48.83 | 47 |
| | **Female** | 126 | 99 | 54.51 | 54 |

Coimbra 1 – The Coimbra Identified Skeletal Collection (CISC);
Coimbra 2 – The 21st Century Identified Skeletal Collection (CEI/XXI).
n1 = number of scanned *ossa coxae;* n2 = number of individuals

## 6.2. Methods

Given the methodological nature of this dissertation, the methods used to fulfil the three main objectives of the thesis can be divided into three sections. The methods are provided only as an overview; a more detailed description is given in the attached publications.

### 6.2.1. Data mining methods applied to visually assessed data

Two independent research teams (Czech and Portuguese) used the dataset (n=941) of visually assessed and scored senescence changes on the pubic symphysis and the auricular surface of ilium. Together, they applied nine different mathematical approaches. The used calculations ranged from less complex to more advanced: the Collapsed regression model, Multi-linear regression, Interval-based model, K nearest neighbours (KNN), Artificial Neural Network (ANN), Decision tree (five classification trees and one

regression tree), M5 tree, and two Probabilistic models (NDE – No Dependence Estimator and AODE – Averaged One-Dependence Estimator). The detailed descriptions of all used models are provided in the study of Kotěrová et al. 2018a (*Appendix A.2*).

The accuracy of individual models was expressed by Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE), defined as (Willmott and Matsuura, 2005; Li and Shi, 2010):

$$MAE = \frac{1}{n}\sum_{j=1}^{n}|y_j - \hat{y_j}|$$

and

$$RMSE = \sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(y_j - \hat{y}_j\right)^2},$$

where $n$ is the sample size, $y_j$ is the estimated age and $\hat{y}_j$ is the actual age.

Moreover, the models were assessed based on whether they are user-friendly, i.e. according to their usability. The performance of classification trees was expressed as a percentage of correctly assigned individuals into specific age classes.

### 6.2.2. Methods used to compare digitized surfaces and to assess the effect on age estimation

First, the acquisition of 3D models of *ossa coxae* was required. The medieval sample (n=29) was scanned with two low-cost scanners, the NextEngine 3D scanner Ultra HD (hereinafter NextEngine) and the HP 3D Structured Light Scanner PRO S2 (hereinafter HP 3D SLS). Unlike the HP 3D SLS, which is based on structured light technology, the NextEngine is based on laser technology. In the case of the HP 3D SLS, during the scanning process (which was conducted under optimal conditions) the whole surface of *os coxae* was scanned several times from various angles. The scanning process was followed by post-processing and the creation of a final 3D model. All these steps were performed in the integrated software David LaserScanner v.3.10.4. Digitization with the NextEngine scanner was done under optimal condition as well. The entire surface of *os coxae*, which was fixed to the rotational device, was scanned and post-processed in the integrated environment of the ScanStudio HD software. All skeletal remains were scanned with

texture, which could be very helpful in the further processing of the 3D models. The final 3D models from both scanners were then simplified to 3 million faces in order to enable easier manipulation with the 3D models. Only pubic symphyseal surfaces were required for further analyses; therefore, they were isolated from the rest of the bone in the MeshLab software (Cignoni et al., 2008).

Additionally, physical casts of five pubic symphyseal surfaces were made and subsequently scanned with the Redlux Profiler, a highly accurate device (Figure 13). Creating the casts was a necessary step, due to the limited space capacity of the Redlux. The scanner consists of linear axes which carry the sensor, and rotary axes which carry the sample. The outputs of Redlux were utilized as reference surfaces for subsequent analyses (resulting surface comparison and age estimation). Data acquisition was provided by A. Kotěrová (HP 3D SLS), L. Friedl (NextEngine), and V. Králík (Redlux).

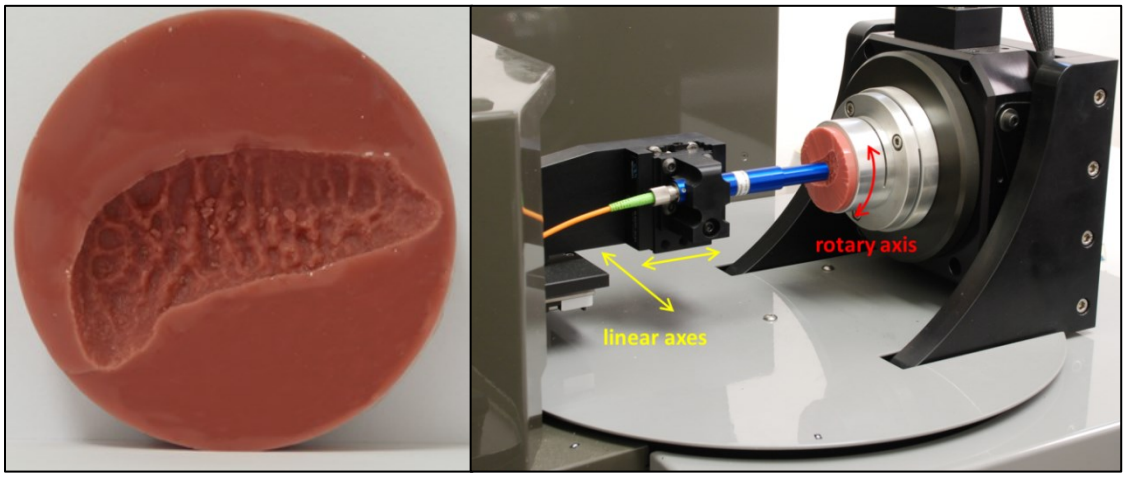


**Figure 13.** Left: Physical cast of symphyseal surface; Right: the Redlux Profiler during the scanning procedure. Photo by Králík Vlastimil. Modified (Kotěrová et al., 2019a).

To compare the resulting 3D surfaces from two low-cost scanners with the reference surfaces, the distribution of deviations was computed and graphically shown using colour-coded maps; the results were evaluated based on descriptive statistics (mean, standard deviation, median, and interquartile range).

The age-at-death estimation was performed with the use of a recent quantitative method proposed by Stoyanova et al. (2015, 2017) in the software 'forAGE'. This computational approach is based on five regression models: two describe the flatness of the pubic symphyseal surface (TPS/BE and SAH score), one characterizes the shape of the

ventral margin (VC) and the last two combine the previous features (TPS/BE+VC and SAH+VC). A paired *t*-test was applied to compare the results of the age estimation derived from two low-cost scanners; a paired one-sample *t*-test was performed to evaluate the differences within the estimations derived from the Redlux scanner and the two low-cost devices. More details can be found in the publication of Kotěrová et al. (2019) attached as *Appendix A.3.*

### 6.2.3. Methodological approach used to validate the aging method of Stoyanova et al. (2015, 2017)

A total of 96 virtual models of male *ossa coxae* from the European samples and 79 virtual models from the Thai sample were selected from the large dataset of the virtual models described above (*Chapter 6.3.*). The selection was limited by several factors. The validated approach (Stoyanova et al., 2015, 2017) was designed with male individuals; therefore, the female individuals were not included in our test sample. Preferably the left-sided *ossa coxae* with intact pubic symphyseal surfaces were used. Finally, the inferior preservation of this part of the pelvic bone and the generally low proportion of individuals under 40 years in the osteological collections did not allow more individuals to be included in the test sample.

Digitization of all *ossa coxae* was performed with the surface scanner HP 3D Structured Light Scanner PRO S2 or with the newer version, the HP 3D Structured Light Scanner PRO S3 (previously known as the David SLS-2 or SLS-3 surface scanner). These two versions of scanners differ slightly only in camera resolution. In the case of the older version PRO S2, the manufacturer stated a resolution of 0.06 mm; the newer version had a resolution of ±0.05 mm. All skeletal samples, except for the one originating from the Khon Kaen collection, were digitized with the HP 3D Structured Light Scanner PRO S2. Only the Thai collection was scanned with the newer version of the scanner (Figure 14). All these steps were performed using the integrated software David LaserScanner v.3.10.4. or v.5.6.0. (HP 3D Scan Software Pro v5). The procedure of scanning with HP 3D SLS is described in the *Methods* section above. For the subsequent analyses, the articulation surfaces of the pubic symphysis were isolated using the MeshLab software, simplified and exported in ply format. All skeletal collections were digitized by A. Kotěrová, except for the Heraklion collection which was digitized by L. Šťovíčková.

**Figure 14.** Surface scanner HP 3D Structured light Scanner PRO S3. Photo by A. Kotěrová.

Age-at-death estimation was performed in the 'forAGE' software developed by Stoyanova et al. (2015, 2017). We have assumed, based on many previous publications (e.g. (Buk et al., 2012; Baccino et al., 2014; Márquez-Grant, 2015)), that pubic symphysis is suitable for estimating the age of approximately 40 years of life, after which the degenerative changes in its surface are too unpredictable. This is also the reason why we provided results for the whole sample, without age restriction, as well as for sub-sets under 40 years and over 40 years. The accuracy of age estimates was expressed as RMSE (defined above), bias and inaccuracy, defined as:

$$Bias = \sum(Estimated\ age - Actual\ age)/N$$

$$Inaccuracy = \sum|Estimated\ age - Actual\ age|/N$$

where *N* is the sample size. A paired *t*-test was used to assess the differences between the actual and estimated ages. More detailed description is provided in the study of Kotěrová et al. (2018b), attached as the *Appendix A.4*.

# 7. RESULTS AND DISCUSSION

## 7.1. Improvement in the evaluation of visually assessed data (Kotěrová et al., 2018a)

Out of the tested approaches, the multi-linear regression model and collapsed regression model performed best; however, the errors were very similar for all models (results for a pooled sample – MAE between 9.7–11 years, and RMSE 12.1–14.2 years). The comparison of model performance is shown separately for males, females and a pooled sample in Table 8.

**Table 8.** Results of MAE and RMSE values for nine mathematical approaches; separately for males, females and a pooled sample. Each model was evaluated according to the ease of application for users. Modified (Kotěrová et al., 2018a).

| | **Males** | | **Females** | | **Pooled** | | **Model is user friendly** |
|---|---|---|---|---|---|---|---|
| | **MAE** | **RMSE** | **MAE** | **RMSE** | **MAE** | **RMSE** | |
| | | | **(years)** | | | | |
| **Collapsed regression model** | 9.3 | 11.6 | 10.4 | 12.7 | 9.9 | 12.2 | Yes |
| **Multi-linear regression** | 9.3 | 11.6 | 10.3 | 12.7 | 9.7 | 12.1 | Yes |
| **Interval-based model** | 10.3 | 12.9 | 12.0 | 14.7 | 10.8 | 13.3 | Yes |
| **K Nearest Neighbours, K=3** | 9.9 | 12.6 | 10.9 | 13.6 | 10.1 | 12.8 | No |
| **Artificial Neural Network** | 11.3 | 14.0 | 11.4 | 14.5 | 11.0 | 14.2 | No |
| **Regression tree** | 10.4 | 12.9 | 11.1 | 14.0 | 10.3 | 12.9 | Yes |
| **M5 tree*** | 9.3 | 11.6 | 10.3 | 12.7 | 9.7 | 12.1 | Yes |
| **NDE**** | 9.8 | 12.8 | 10.9 | 13.9 | 10.4 | 13.3 | No |
| **AODE**** | 9.8 | 12.7 | 10.6 | 13.6 | 10.2 | 13.2 | No |

MAE - mean absolute error; RMSE - Root Mean Square Error; NDE - No Dependence Estimator; AODE - Averaged One-Dependence Estimator,
* M5 tree collapsed into one leaf, thus it becomes equivalent to multi-linear regression
** Reference sample age distribution used as the prior distribution

Among the classification trees, the highest accuracy (72%) was reached with the single indicator, the pubic symphysis, in classification up to three age intervals (<30 years, 30–40 years, and >40 years). The lowest, on the other hand, was attained with the tree based on both indicators, in classification up to 10-years intervals (31%). Results are shown in Table 9. Complete results are included in the study Kotěrová et al., 2018a, which is attached as *Appendix A.2.*

**Table 9.** Comparison of classification tree models, which assign the examined person into a specific class, pooled sample, depth 4 (Kotěrová et al., 2018a).

| | **Tree 1** | **Tree 2** | **Tree 3** | **Tree 4** | **Tree 5** |
|---|---|---|---|---|---|
| **Mean accuracy (%)** | 69.0 | 68.6 | 30.7 | 72.3 | 58.6 |

Tree 1 – both indicators, age interval <30, 30-60, >60 years; Tree 2 - both indicators, age interval <30, 30-50, >50 years; Tree 3 - both indicators, age interval 10-years intervals; Tree 4 – PUSx indicator, age interval <30, 30-40, >40 years; Tree 5 – SSPIx indicator, age interval <30, 30-40, 40-50, >50 years.

### *Discussion*

The study of Kotěrová et al. (2018a) aims to improve the accuracy of age-at-death estimation by applying various mathematical approaches. However, the results showed that no matter how sophisticated the approaches that are applied to the data may be, none of them led to significant improvement in age-at-death estimation from the skeleton. These results are in concordance with the studies of Schmitt et al., Buk et al. and Martins et al., who used the same or similar dataset and applied different data mining methods (Schmitt et al., 2002; Buk et al., 2012; Martins et al., 2012). As has been shown in our study and many studies before, every effort has been made for decades to minimize estimation errors, with an emphasis on both the anatomical structures used and the statistical treatment of the data. Nevertheless, the age estimation situation is still equally unsatisfactory.

We assume that the explanations of this dismal situation regarding age estimation methods may be as follows:

a) The subjective nature of age-related changes evaluation, since most of the aging methods are based on the visual assessment of age indicators (e.g. (Brooks and Suchey, 1990; Buckberry and Chamberlain, 2002; Rissech et al., 2006; Rougé-Maillart et al., 2009))

b) The assumption that a single indicator can reflect the whole adult lifespan (e.g. (İşcan et al., 1984b; Calce, 2012))
c) The population variability in the aging process (e.g. (Schmitt, 2004; Mays, 2014; Navega et al., 2018b))
d) The effect of external factors on the rate of the senescence process (e.g. (Campanacho et al., 2012; Merritt, 2015; Wescott and Drew, 2015))
e) The extreme variability and unpredictability of the aging process.

We believe that the visual assessment of senescence changes remains the limiting factor of accuracy. The possible direction of improvement may lie in the substitution of the subjective visual approaches of morphological change evaluation by more objective mathematical surface quantification. We consider our study as a springboard for further research which should focus on the objectification of data acquisition.

## 7.2. The influence of various scanning devices on captured surfaces and on age-at-death assessment (Kotěrová et al., 2019)

A comparison of five 3D pubic symphyseal surfaces captured with the HP 3S SLS and the NextEngine, with reference surfaces acquired by the Redlux scanner, showed small deviations, particularly in areas of depressions and protrusions. The deviations are visualized in the form of colour-coded maps using an example of two samples in Figure 15, where the red scale indicates that the surface to be compared is above the reference surface and blue scale indicates the opposite. The differences between the reference and compared surfaces are more noticeable in the output from the NextEngine scanner.

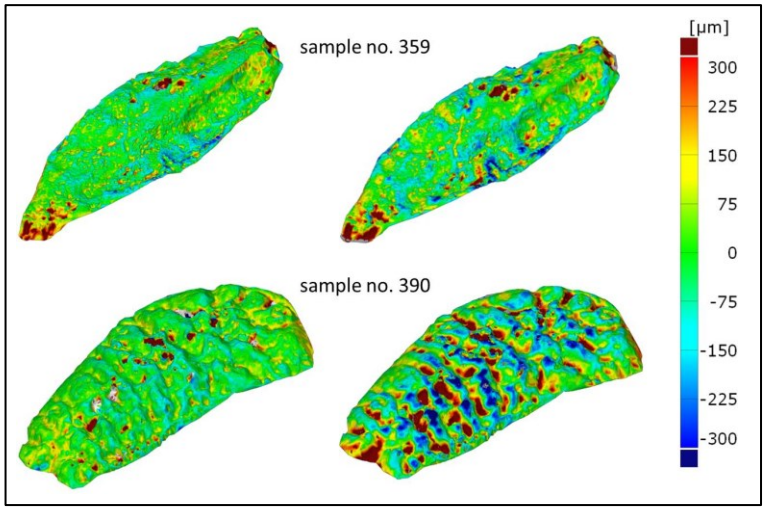


**Figure 15.** Colour-coded map illustrating deviations between the reference and the compared surface (left: HP 3D SLS; right NextEngine) shown on two samples. Positive scale (in red) - the compared surface is above the reference surface; Negative scale (in blue) - the compared surface is below the reference surface (Kotěrová et al., 2019).

Figure 16 shows the graphic comparison of all five compared surfaces (in blue – the reference surface vs HP 3D SLS; in red – reference surface vs NextEngine) in the form of box plots; Table 10 summarizes the descriptive statistics. It could be seen that the deviations from the reference surfaces are slightly higher in the case of the NextEngine. The interquartile range and standard deviation are both higher in all samples (except for one specimen).

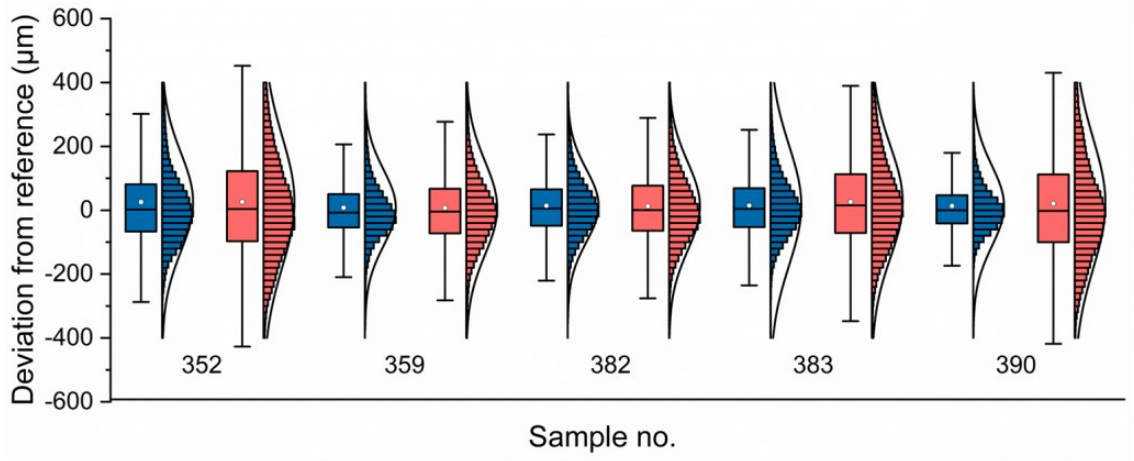


**Figure 16.** Deviation from the reference sample (Redlux scan) for the HP 3D SLS (blue) and NextEngine laser scanner (red). On each box, the central mark is the median, the white point is the mean value for normal distribution µ, the edges of the box are the 25th and 75$^{th}$ percentiles, the whiskers extend to the most extreme data points not considered outliers (Kotěrová et al., 2019).

**Table 10.** Results of quantitative comparison among measured samples (n=5; HP 3D SLS vs. Reference sample and NextEngine vs. Reference sample). Modified (Kotěrová et al., 2019).

| Specimen | Scanner | Mean value for normal distribution | Median | Standard deviation ($\mu m$) | Interquartile span | Percentage difference *IQR* (%) |
|---|---|---|---|---|---|---|
| 352 | HP 3D | 26 | 2 | 154 | 147 | +50 |
| | NextE | 26 | 4 | 204 | 220 | |
| 359 | HP 3D | 8 | -8 | 118 | 104 | +35 |
| | NextE | 7 | -4 | 146 | 140 | |
| 382 | HP 3D | 14 | 6 | 113 | 114 | +24 |
| | NextE | 13 | 0 | 138 | 141 | |
| 383 | HP 3D | 15 | 5 | 187 | 122 | +51 |
| | NextE | 26 | 15 | 174 | 184 | |
| 390 | HP 3D | 14 | 0 | 126 | 88 | +141 |
| | NextE | 21 | 2 | 202 | 212 | |

HP 3D – HP 3D Structured light scanner; NextE – NextEngine scanner

The results of the age-at-death analyses based on 3D models acquired with the HP 3D SLS and NextEngine scanners showed that, although the surfaces varied slightly, this did not significantly affect the age analyses performed. No significant differences were observed among the estimated ages (with one exception) based on the models from HP 3D SLS and NextEngine (Table 11).

| Regression model | *p*-value |
|---|---|
| **TPS/BE**[1] | 0.01 |
| **SAH**[2] | 0.07 |
| **VC**[3] | 0.10 |
| **TPS/BE+VC**[4] | 0.27 |
| **SAH+VC**[5] | 0.41 |

[1] Thin plate spline/Bending energy, [2] Slice and Algee-Hewitt score, [3] Ventral curvature, [4] combination of TPS/BE and VC (multivariate model), [5] combination of SAH score and VC (multivariate model).
*p*-values are shown for a two-tailed test.

**Table 11.** Results of a paired *t*-test between the estimated ages based on outputs from the two low-cost scanners (n=29). Age estimation was made according to the method of Stoyanova et al., 2017 (Kotěrová et al., 2019).

The references surfaces from the Redlux were subjected to age analysis as well. Differences between HP 3D SLS and RedLux for the SAH model, and between NextEngine and Redlux in the SAH and SAH+VC models were detected. In all other cases the null hypothesis was not rejected. Complete results can be found in Kotěrová et al. (2019).

***Discussion***

With the growing interest in scanning technologies and their integration into common practice (e.g. estimation of biological profiles) in anthropological workplaces and forensic institutions, questions have arisen about the comparability of outputs from different devices and their impact on subsequent biological profile analyses. These questions first arose from Villa et al. (Villa et al., 2015b). They looked at the potential differences in the surface recording of the two pelvic articular surfaces (the pubic symphysis and the auricular area) among the three different laser scanners and their impact on the surface quantification approaches for age assessment proposed by Villa and colleagues (Villa et al., 2015a). Their results indicated a comparable overall anatomical shape, but the curvature values showed systematic bias among the three tested scanners, although the curvature values changed similarly with increasing phase or score.

In the study of Kotěrová and co-workers, the slightly modified original inquiry of Villa et al. (2015a) was examined. Specifically, we were interested in the comparability of outputs from scanners representing various technologies (laser and structured light) and their influence on age estimation, as well as on sex assessment (although the latter was not presented in the dissertation thesis, the results are part of the publication). Given that research institutions around the world are equipped with different types of scanning devices, these results are of utmost importance and our results could have a positive impact on data sharing among researchers and institutions. Although the NextEngine captured less detail and the information loss was the largest, the differences, according to our results, were relatively small with expected negligible effect on further analyses. Our assumptions were confirmed in that it seems that the type of surface scanner and its various resolution capabilities do not affect the estimation of age-related changes, at least not the changes quantified by the method of Stoyanova and colleagues (2015, 2017). Given these results, the method of Stoyanova et al. can be applied to data obtained by other surface scanners.

However, other scanners and scanning technologies have to be considered and tested (Villa et al., 2015b; Kotěrová et al., 2019). Therefore, in the near future, it is necessary to repeat the tests and quantify the influence of various scanning devices on the results of age estimation and other biological profile analyses. We assume that age estimation based on the mathematical quantification of surface changes will be the most susceptible to differences in captured surface by different devices since even subtle surface age-related changes are assessed.

## 7.3. Reproducibility and validity of the Stoyanova et al. aging method (2015, 2017) in geographically different populations (Kotěrová et al., 2018b)

The results of the reproducibility and validity of recent computational methods will be presented separately, first for the pooled European samples and then for the Thai sample.

In our pooled European sample, the validated method performed with unacceptably high values of RMSE, bias and inaccuracy (e.g. RMSE values ranged between 18.35 to 22.25 years). In the sample over 40 years, the method performed the worst. Conversely, the results that were obtained in our sub-sample of under 40 years seem very promising. In this sample, the results of RMSE were between 5.93 and 7.48 years (apart from VC), and bias and inaccuracy ranged between −1.77 (TPS/BE+VC) and 5.90 years (VC), and 4.47 (TPS/BE) and 7.07 (VC) years, respectively. In the sample up to 40 years, the differences between actual and estimated ages were not significant (with the exception of VC). There was systematic underestimation of actual age in all the age categories of our sample. Table 12 compares the results reported in the study of Stoyanova et al. and that of Kotěrová and colleagues (2018b). More detailed information is provided in the study of Kotěrová et al. 2018b (*Appendix A. 4*).

**Table 12.** Comparison of the values of RMSE, bias and inaccuracy (in years) in the original study of Stoyanova et al. (2017) and in the study of Kotěrová et al. (2018b). Results are presented for the whole dataset and the sample of up to 40 years (Kotěrová et al., 2018b).

| | | Whole dataset | | 20–40 years | |
|---|---|---|---|---|---|
| **Regression model** | | **Kotěrová et al. (2018b)** | **Stoyanova et al. (2017)** | **Kotěrová et al. (2018b)** | **Stoyanova et al. (2017)** |
| **TPS/BE** | RMSE | 22.09 | 16.38 | 5.93 | – |
| | Bias | −15.54 | −2.51 | −1.64 | 8.47 |
| | Inaccuracy | 16.75 | 12.58 | 4.47 | 9.08 |
| **SAH** | RMSE | 20.91 | 14.15 | 7.34 | – |
| | Bias | −13.58 | −1.96 | −0.78 | 5.71 |
| | Inaccuracy | 15.70 | 10.81 | 5.75 | 8.10 |
| **VC** | RMSE | 18.35 | 16.55 | 9.49 | – |
| | Bias | −8.43 | −2.73 | 5.90 | 5.84 |
| | Inaccuracy | 14.15 | 12.86 | 7.07 | 9.89 |
| **TPS/BE+VC** | RMSE | 22.25 | 15.07 | 6.37 | – |
| | Bias | −15.67 | −2.21 | − 1.77 | 7.76 |
| | Inaccuracy | 16.96 | 11.39 | 4.79 | 8.93 |
| **SAH+VC** | RMSE | 21.8 | 13.68 | 7.48 | – |
| | Bias | − 13.67 | −1.82 | −0.70 | 5.76 |
| | Inaccuracy | 15.99 | 10.79 | 6.13 | 8.36 |

TPS/BE − thin plate spline/bending energy; SAH − Slice and Algee-Hewitt score; VC − ventral curvature.
RMSE for age category 20–40 years in Stoyanova et al. (2017) was not calculated

The results presented for the Thai sample were not included in the publication of Kotěrová et al. (2018b) and they are not yet published.

Again, the resulting RMSE values for both the whole sample and the sample over 40 years are shown to be unacceptably high (ranging from 15.71 to 20.09 years and 20.44 to 26.23 years, respectively). For the sub-sample of up to 40 years, the RMSE values were between 6.35 and 7.99 years (except for VC). The constant underestimation of actual age is obvious, except for the SAH, VC and SAH+VC models in the up to 40 years sample, where age overestimation was observed. The differences between actual and estimated age were statistically insignificant only for the SAH, TPS/BE+VC and SAH+VC regressions models in the age category under 40 years. The results are summarized in Table 13.

**Table 13.** Results of RMSE, Bias, Inaccuracy and paired t test for the whole sample from the Thai population; sample under 40 years and over 40 years.

| Regression model | Age category | RMSE | Bias (years) | Inaccuracy | *p* value | |
|---|---|---|---|---|---|---|
| | | | | | Uncorrected | Corrected |
| **TPS/BE** | All data | 19.34 | −13.70 | 14.83 | <0.001 | <0.001 |
| | ≤ 40 years | 6.35 | −3.28 | 5.50 | 0.001 | 0.005 |
| | > 40 years | 25.57 | −22.42 | 22.65 | <0.001 | <0.001 |
| **SAH** | All data | 15.71 | −8.26 | 11.40 | <0.001 | <0.001 |
| | ≤ 40 years | 6.53 | 0.27 | 5.31 | 0.81 | 1* |
| | > 40 years | 20.44 | −15.40 | 16.50 | <0.001 | <0.001 |
| **VC** | All data | 18.14 | −7.16 | 13.46 | <0.001 | <0.001 |
| | ≤ 40 years | 9.84 | 5.37 | 7.52 | <0.001 | 0.002 |
| | > 40 years | 22.88 | −17.65 | 18.44 | <0.001 | <0.001 |
| **TPS/BE+VC** | All data | 20.09 | −13.16 | 15.19 | <0.001 | <0.001 |
| | ≤ 40 years | 7.99 | −2.33 | 6.68 | 0.08 | 0.40* |
| | > 40 years | 26.23 | −22.22 | 22.32 | <0.001 | <0.001 |
| **SAH+VC** | All data | 16.82 | −8.51 | 11.98 | <0.001 | <0.001 |
| | ≤ 40 years | 7.44 | 0.61 | 5.82 | 0.63 | 1* |
| | > 40 years | 21.75 | −16.14 | 17.14 | <0.001 | <0.001 |

*p* values were corrected with the Bonferroni correction. Non-significant differences between true and estimated ages are marked with asterisks
TPS/BE − thin plate spline/bending energy; SAH − Slice and Algee-Hewitt score; VC − ventral curvature

### *Discussion*

Given that fact that the enhanced computational approach developed by Stoyanova et al. (Stoyanova et al., 2015, 2017) was proposed quite recently, there are not many studies that have evaluated the validity and applicability of this computational approach. Our team was the first to publish the validation results of this approach on a pooled sample of European populations. Later, another validation was performed on a European sample, but presented only as a poster (Johnson and Bethard, 2019). In 2019, a publication was published using a South African population (Joubert et al., 2019) and, finally, for the first time, the results on a South Asian population are presented in this thesis.

Our results are consistent with those made by other researchers who aimed to validate the method in their sample; all validation studies reached very similar conclusions. Stoyanova and colleagues compared the accuracy (2× RMSE value) of her results (Stoyanova et al., 2017) to the widely used method of Suchey and Brooks (Brooks and Suchey, 1990) and found her age intervals generally smaller (intervals ranged from 27 to

32 years). Nevertheless, this is not the case for the authors testing her approach. Kotěrová et al. (2018b) report that the intervals ranged from 36 to 44 years for the mix of European samples, which is approximately comparable to the performance of the Suchey-Brooks method. Slightly better results were reached for the Thai population (age interval ranging from 30–40 years). Johnson and Bethard (2019) reached extremely broad age intervals, ranging from 60 to 82 years in the modern Coimbra collection. Joubert et al. (2019) did not provide RMSE values from which age intervals are calculated; however, they reported very low correlations with true age in their South African samples.

Such unsatisfactory results can be caused by several factors. Firstly, there are different aging rates of the pubic symphysis among these populations. Individuals from various geographical regions were exposed to different internal and external factors, which may affect the rate of degeneration and thus the performance of the proposed method. The population specificity of the aging methods was noted (e.g. (Mays, 2014; Savall et al., 2016; Hisham et al., 2019)), although some researchers did not reach such conclusions (e.g. (Sakaue, 2006; Kimmerle et al., 2008)).

Secondly, the pubic symphysis was proved to be useful as an age indicator only up to approximately 40 years (Lovejoy et al., 1995; Schmitt et al., 2002; Sakaue, 2006; Buk et al., 2012; Gocha et al., 2015). The samples of Joubert et al. (2019) and Johnson and Bethard (2019), which, moreover, were very small (n=36), were predominantly represented by old individuals. The sample of Joubert et al. (n=184) was strongly underrepresented by younger individuals under 45 years (n=28) and the mean age of the European sample used by Johnson and Bethard was 66.8 years. This could be one of the reasons why they achieved such poor results. We divided our mixed European and Thai samples (Kotěrová et al., 2018b; Kotěrová unpublished) into the samples up to 40 years and over 40 years with the assumption that a sample limited to 40 years of age would be estimated more accurately. As these assumptions have been confirmed, the results are promising but still not ideal. The lowest RMSE was 5.93 years, giving an age interval of 12 years; however, that was only in the sample up to 40 years. Therefore, Joubert et al. (2019), as well as our team, advocate for the use of multi-indicator methods to encompass more variation with regard to their biological possibilities (Kotěrová et al., 2018b).

Thirdly, the input data are isolated surfaces of the pubic symphysis, thus it is necessary to manually isolate them, which is a possible source of errors influencing the analyses. However, this question was raised by Joubert et al. (2019), who stated that the

repeatability is very strong in comparison to traditional phase-based methods. The excellent reproducibility among the five observed was confirmed by the extended original research team (Kim et al., 2018).

Fourthly, it can be argued that 3D representations of skeletal samples acquired with various scanning devices may not be comparable (Joubert et al., 2019; Kotěrová et al., 2019). While the NextEngine was employed in the original study (as well as in the work of Johnson and Bethard, 2019), the Artec Space Spider scanner was used by Joubert et al. (2019), and the HP 3D SLS was used by our team. This issue was addressed in our previous research and the results showed that various scanning technology does not affect age estimation.

Undoubtedly, approaches such as the one suggested by Stoyanova et al. have a great potential to deal with some problems accompanying age-at-death estimation. Although the performance of Stoyanova's method is quite comparable to conventional methods and did not bring higher accuracy, its great potential lies in reducing subjectivity. The increase of accuracy is needed to make the method of Stoyanova et al. more acceptable for widespread use.

The validation study has shown that the limiting age of the pubic symphysis is the chronological age of approximately 40 years. It is therefore not appropriate to use methods based on that structure alone to estimate the age-at-death of unknown skeletal remains. It seems that a kind of a hierarchical approach is needed in which multiple age indicators would be used gradually, not at the same time, as this would lead to unclear results. This means that after reaching the optimum age estimation with the pubic symphysis, the auricular surface could be used next. Once this indicator had also reached its optimum, senescence changes in the acetabulum could be used. As there are still some theoretical assumptions for this procedure, it must be verified and tested.

# 8. DIRECTION FOR FUTURE RESEARCH

Further research will be directed towards the design of our own age-at-death estimation approach. With regards to the results of this doctoral thesis, the aim will be to eliminate some of the pitfalls of age estimates that are preventing more reliable and accurate age estimation.

Firstly, we would like to achieve this by eliminating the subjective evaluation of senescence changes by its replacement with a fully quantitative approach analysing three-dimensional surface data. The creation of an extensive digital database of 3D models of *ossa coxa* would be an essential step for further research. Secondly, the multi-population composition of the database is the cornerstone for the development of robust approaches that promise wider applicability than single-population-based methods. Thirdly, given that the database consists of the entire surfaces of the *ossa coxae,* we now have at our disposal all three articulation surfaces present on the *os coxae* (the pubic symphyseal surface, the auricular surface of ilium, and the surface of the acetabulum). We aim to use these joint surfaces to propose our own approach to assess age-at-death. However, the articulation surfaces will not be used to estimate age together at the same time, but rather they will be chronologically combined on the basis of predetermined information in which the period of the adult life in the articular area is the most appropriate for age estimation.

At the time of finalizing this doctoral thesis, we have been developing our own approaches on the pubic symphyseal surface. They aim to extract specific signs or patterns that will help to distinguish between individuals over 40 years and classify individuals under 40 years more specifically. One of the approaches that are being tested is the segmentation of elevated areas with the use of the polynomial function of the 4th degree, which helps transform the shape of an original surface into a form that allows better analysis of its local irregularities. Another approach that we are currently testing is ‘deeper’ analysis, which computes the discrete Gauss curvature in each point of the mesh (i.e. the surface of the symphyseal articular area). We expect that both approaches could capture age-related changes of the surface that cannot be captured by visual methods. At the very least, we assume they will objectify the capturing of surface changes.

Once these analyses are finished and the best one is optimized, the other two surfaces of the *os coxae* will be subjected to detailed analysis as well. Based on data mining methods, the outputs for all three age indicators will be chronologically combined

and incorporated into user-friendly software that will be developed for automatic and faster age estimation (once 3D representation of the surface is made), without subjectivity or the need for special experience.

What more could be done to more accurately estimate age? If the proposed fully quantitative mathematical approaches based on a chronological combination of age indicators (on which we are working) do not lead to an increase in the accuracy and reliability of the methods, then all the possibilities to achieve this will probably have been exhausted. It would have to be admitted that for now the aging process is simply too variable and unpredictable.

# 9. CONCLUSIONS

Three main objectives were set out in this dissertation thesis. Firstly we aimed to improve age-at-death estimation based on traditional visual assessment of age-related changes on the pubic symphysis and the auricular area with the use of various mathematical approaches. Nevertheless, our results suggest that a broad age-range reduction is not possible. We have found that even with various sophisticated mathematical approaches, the accuracy of methods was not increased, but remained approximately the same as that provided by the previous methods. We believe this problem will persist as long as the subjective methods based on visual assessment continue to be used to estimate age. The only solution is the substitution of the visual evaluation of senescence patterns with virtual anthropology, scanning technologies and mathematical quantification of surface age-related changes.

The second objective was to assess and compare the outputs from various surface scanners and to find out whether the potential differences in the captured surfaces could have an effect on further analyses (e.g. age-at-death estimation based on quantitative surface analysis). According to our results, even though small differences in captured surfaces were detected and the NextEngine turned out to be a less precise scanner (capturing less detail than the HP 3D SLS scanner), it did not significantly affect age estimation. Researchers may use 3D data that are publicly available or provided by other researchers or institutions, knowing that they will not affect their further analyses even though they were acquired with various scanning devices (under optimal scanning settings). However, such validations are also needed for other scanning devices, as well as for other methods of biological profile estimation.

As a necessary step toward method objectification and a prerequisite for further research, an extensive, multi-population identified database of 3D models of articular surfaces of *ossa coxae* was created. Without this database, the methods of age estimation based on the objective observation of changes on joint relief could not be tested and new approaches could not be proposed. The database consists of more than 1,200 scans of *ossa coxae* of identified adult males and the female skeletal remains of four European and one Asian osteological collections. Such a multi-population database establishes a solid basis for covering more variability in the aging process and for more objective data evaluation.

The third objective was the validation of the recent computational approach proposed by Stoyanova et al. (2015, 2017) which was performed using the created database, first on a mix of European samples and then on an Asian sample, which was digitized later. Earlier findings known from literature were confirmed even by quantitative analyses. Specifically, the pubic symphysis after 40 years of life undergoes overly variable degenerative changes that result in overly inaccurate estimates (RMSE = 15.7–22.35 years) with overly wide age ranges (2× RMSE) that then cannot be used for age estimation (VC values, whose performance was the worst, were omitted). On the other hand, the method of Stoyanova et al. performed quite well in our European and Asian samples limited to 40 years with the lower RMSE values of 5.93 years and the highest of 7.99 years (apart from the VC values), which is even better than in the original study.

The dissertation is concluded with very clear prospects for future research. The collected metapopulation data allow for the testing of other age indicators, such as the auricular surface of ilium or the acetabulum.

# SOUHRN (CZECH SUMMARY)

Jedním z prioritních cílů forenzních antropologů a bioarcheologů je odhad biologického profilu neznámých kosterních pozůstatků. Biologický profil je tvořen čtyřmi hlavními osteologickými ukazateli, někdy nazývanými jako „velká čtyřka“, kterými jsou odhad pohlaví, věku, biogeografického původu (známého také jako odhad populační afinity či odhad původu) a výšky postavy (Cattaneo, 2007; Ellingham and Adserias-Garriga, 2019). Mezi těmito parametry je právě odhad věku dožití klíčovým parametrem, protože ve forenzní antropologii může významně zúžit okruh potenciálních identifikovaných jedinců. V bioarcheologii pak poskytuje cenné informace o demografickém profilu populací. Zároveň je ale parametrem nejproblematičtějším (Adams and Byrd, 2014; Algee-Hewitt, 2017; DeWitte, 2017), zatíženým nepřesnými a nespolehlivými odhady.

Metody odhadu věku dožití jsou založeny na korelaci mezi biologickým a chronologickým (kalendářním) věkem. U nedospělých jedinců, jejichž skelet stále prochází dobře predikovatelnými vývojovými a růstovými změnami, je tento vztah velmi úzký a dovoluje přesný i spolehlivý odhad věku (Franklin, 2010). Oproti tomu, u plně dospělých jedinců s ukončeným růstem a vývojem, makroskopické metody odhadu věku tedy mohou být založeny pouze na pozorování degenerativních senescenčních procesů na kloubních plochách, např. stydké sponě a aurikulární oblasti pánevní kosti, kyčelní jamce či sternálních koncích žeber (Christensen et al., 2014b; Langley et al., 2017). Proces stárnutí je velmi komplexní. Degenerativní morfologické změny jsou s narůstajícím věkem čím dál více ovlivňovány řadou faktorů, a to mnohem více, než vývoj a růst u nedospělých jedinců. Vztah mezi chronologickým a biologickým věkem se tak rozvolňuje, což v důsledku vede k nepřesným odhadům. Proces stárnutí je variabilní jednak mezi jedinci stejné populace, či určité ekogeografické skupiny, ale i mezi těmito skupinami. Populační variabilita věkových změn je známá. Vznikající metody jsou většinou populačně specifické a netýkají se pouze odhadu věku. Nerespektování populační specificity metod může vést k velmi zkresleným výsledkům (Mays, 2014; Kotěrová et al., 2017; Navega et al., 2018b). Odhad věku je však limitován mnoha dalšími faktory, mezi kterými zmíníme nevhodně zvolené statistické metody, či mylnou představou, že jediný věkový indikátor může postihnout celý život dospělého jedince. Výsledkem této situace je možnost odhadovat věk spolehlivě a přesně současně, je-li užito tří širokých věkových intervalů (např. do 30 let, 30

až 60 let a nad 60 let, jak navrhují Buk a kolektiv (2012), Falys a Lewis (2011), či Adserias-Garriga and Wilson-Taylor (2019)). V soudně lékařské praxi nejsou tyto široké věkové intervaly vhodné pro zúžení okruhu potenciálních identifikovaných osob a je nutno jej zpřesnit. Řešení mohou přinést jednak sofistikované matematicko-statistické přístupy (např. Bayesovské přístupy, rozhodovací stromy, neuronové sítě), ale také metody virtuální antropologie, které slibují objektivizaci hodnocení senescenčních změn kloubních povrchů.

Na základě současného stavu poznání o odhadu věku dožití byly v předkládané disertační práci vytyčeny tři hlavní cíle. Prvním z cílů byla snaha o zpřesnění klasifikace věku klasickými vizuálními metodami užitím různých matematických přístupů. Dalším cílem bylo porovnat povrchy 3D modelů kloubní plochy stydké spony zachycené různými povrchovými skenery a posoudit vliv případných odchylek na analýzu odhadu věku kvantitativní metodou a možnost sdílení referenčních dat. Posledním stěžejním cílem byla validace recentního kvantitativního přístupu (Stoyanova et al., 2015, 2017) odhadu věku podle pubické symfýzy.

Naším prvním cílem bylo posoudit možnosti přesnějšího a spolehlivějšího odhadu věku dospělých jedinců vizuálně hodnocenými postupy za využití různých sofistikovaných matematicko-statistických metod. Naším záměrem bylo zúžit úsek dospělého života na více než tři široké věkové intervaly navržené Bukem a kolektivem (Buk et al., 2012). Využili jsme multi-populační vzorek (n=941) dvou pánevních artikulačních plošek (pubické symfýzy a aurikulární oblasti pánevní kosti) ohodnocených A. Schmitt (Schmitt, 2001, 2005; Schmitt et al., 2002). Dva nezávislé týmy výzkumníků, český a portugalský, pak na tato data aplikovaly devět různých matematických přístupů ve snaze dosáhnout co nejpřesnějšího odhadu věku dožití. Testovanými přístupy a klasifikačními technikami byly: regresní analýza kumulující proměnné (Collapsed regression model), multi-lineární regrese, model založený na intervalech (Interval-based model), klasifikace podle nejbližších sousedů (K nearest neighbours), umělá neuronová síť (Artificial Neural Network), rozhodovací strom (Decision tree), metoda M5 tree, a dva probabilistické modely (NDE – No Dependence Estimator and AODE – Averaged One-Dependence Estimator). Výsledky byly vyjádřeny pomocí hodnot MAE a RMSE zvlášť pro ženy, muže a pro smíšený soubor. Mezi těmito přístupy se jako nejlepší ukázaly Multi-lineární regrese a regresní kumulativní model, ačkoli hodnoty RMSE (12,1–14,2 let) a MAE (9,7–11 let) byly velmi obdobné. Mezi klasifikačními stromy bylo nejvyšší přesnosti dosaženo na

základě jednoho indikátoru do tří věkových skupin (<30 let, 30–40 let, a >40 let). Ke zpřesnění klasifikace za využití různých matematických přístupů však nedošlo. Široké věkové intervaly jsou nutností při vizuálním hodnocení věku dožití dospělých jedinců. Hlavním limitujícím faktorem je proto subjektivní podstata hodnocení věkových změn na kloubních plochách. Cestu ke zlepšení vidíme v objektivizaci získání vstupních proměnných a následné hodnocení dat pomocí nástrojů virtuální antropologie. Tyto výsledky byly publikovány ve studii Kotěrové et al. (2018a).

Druhým cílem této práce bylo porovnat povrchy pubické symfýzy digitalizované různými skenovacími zařízeními a zhodnotit vliv případných odchylek zachycených povrchů na následné analýzy odhadu věku. Pro tuto výzkumnou část jsme různými skenery naskenovali 29 pánevních kostí z české středověké populace. Digitalizace byla provedena jednak laserovým skenerem NextEngine, dále skenerem HP 3D SLS založeným na strukturovaném světle a malý vzorek (n=5) byl navíc naskenován skenerem Redlux Profiler s vysokým rozlišením, jehož výstupy byly považovány jako referenční. Rozdíly mezi porovnávanými povrchy ze skenerů NextEngine a HP 3D SLS s referenčními povrchy byly vyjádřeny výpočtem distribuce odchylek, dále graficky na barevných mapách a výsledky byly zhodnoceny pomocí deskriptivní statistiky. Odhad věku byl proveden kvantitativní metodou navrženou Stoyanovou a kolektivem (2015, 2017), která je založena na pěti regresních modelech inkorporovaných do uživatelsky jednoduchého softwaru „forAGE“ (Stoyanova et al., 2017). Vstupními daty jsou 3D modely izolovaných ploch pubické symfýzy. Dva z regresních modelů hodnotí „plochost“ těchto povrchů (TPS/BE a SAH skóre) a další model charakterizuje zakřivení ventrálního okraje plošky (VC). Zbývající dva modely jsou kombinací modelů předešlých (TPS/BE + VC a SAH + VC). Výsledky srovnání digitalizovaných povrchů odhalují malé odchylky od referenčních povrchů, které jsou patrné především v oblastech depresí a protruzí. Tyto rozdíly byly výraznější mezi referencí a výstupy ze skeneru NextEngine. Analýza odhadu věku dožití na základě kloubních ploch pubické symfýzy získaných ze skenerů NextEngine a HP 3D SLS neodhalila žádné významné rozdíly mezi získanými odhady věku. Jedinou výjimkou byl odhad věku na základě modelu TPS/BE. Naše výsledky ukazují, že odlišné typy povrchových skenerů a jejich rozdílné rozlišovací schopnosti neovlivňují analýzu odhadu věku, i když se získané povrchy od sebe lehce liší. Nicméně podobné analýzy a testy je nutné provést i pro další skenovací zařízení, stejně tak, jako pro ostatní analýzy

biologického profilu. Kompletní výsledky jsou uvedeny v publikaci Kotěrové a kolektivu (2019).

Posledním úkolem, který jsme si vytyčili, bylo validovat kvantitativní metodu Stoyanové a kolektivu (2015, 2017) v geograficky odlišných populacích. K tomuto účelu bylo z rozsáhlé multi-populační identifikované databáze 3D modelů artikulačních povrchů pánevní kosti, kterou jsme vytvořili, vybráno 96 povrchových modelů mužských pánevních kostí ze smíšeného evropského souboru a 79 modelů z thajského souboru tak, aby distribuce věku byla rovnoměrná. Digitalizovány byly povrchovým skenerem HP 3D SLS. Před vlastní validací metody v softwaru „forAGE“ byly všechny artikulační plošky pubické symfýzy izolovány v programu MeshLab. Výsledky byly vyjádřeny pomocí hodnot vyjadřujících průměrné nadhodnocení a podhodnocení skutečného věku jedince (bias), dále pomocí hodnot vyjadřujících průměrnou absolutní chybu odhadu (inaccuracy) a hodnotami RMSE. Na základě mnoha předchozích publikací jsme předpokládali, že věkově vázané změny na pubické symfýze jsou pro odhad věku vhodné pouze přibližně d 40 let (např. (Baccino et al., 2014; Márquez-Grant, 2015)). Proto jsme všechny výsledky uváděli nejen pro celý soubor, ale také pro omezený soubor do 40 let a nad 40 let. Naše výsledky jasně ukázaly, jak pro smíšený evropský soubor, tak pro thajský soubor, že metoda Stoyanové a kolektivu (Stayanova et al., 2015, 2017) je zatížena nepřijatelně vysokou chybou odhadu, pokud je aplikována na celý, věkově neomezený, soubor. Chyba se ve smíšeném evropském souboru pohybovala v hodnotách RMSE v rozmezí od 18,35 do 22,25 let, v thajském souboru o něco níže, v rozmezí od 15,71 do 20,09 let. Chyba odhadu v souboru omezeném nad 40 let byla ještě vyšší. Nicméně mnohem nižních chyb bylo dosaženo v souborech do 40 let. V evropském souboru byla nejnižší chyba 5,93 let dosažena modelem TPS/BE, rovněž tak v thajském souboru, kde byla nepatrně vyšší (6,35 let). V této výzkumné části se nám podařilo potvrdit, že pubická symfýza není vhodná pro odhad věku u jedinců starších 40 let, chyba odhadu je pak nepřijatelně velká. Výsledky ve věkově omezeném souboru do 40 let jsou sice slibné, nicméně chyba odhadu 5,93 let představuje věkový interval 12 let, což stále ještě není uspokojivý výsledek. Kompletní výsledky validace v evropském souboru jsou publikovány ve studii Kotěrová a kolektivu (2018b), výsledky validace v thajském souboru jsou dosud nepublikované a prvně uvedené v této disertační práci.

Některé výzkumné části této práce byly podmíněny vytvořením multi-populační identifikované databáze 3D modelů artikulačních povrchů pánevní kosti. Bez této databáze

by nemohly být metody odhadu věku založené na kvantitativních analýzách validovány, ale ani navrhovány. Z této databáze budeme vycházet i v dalším našem výzkumu.

# BIBLIOGRAPHY


Abegg C, Desideri J. 2018. A probable case of multiple myeloma in a female individual from the Simon Identified Skeletal Collection (late 19th–early 20th century, Vaud, Switzerland). Int J Paleopathol 21:158–165.

Acsádi G, Nemeskéri I. 1970. History of human life span and mortality. Budapest: Akademiai Kiado.

Adams B, Byrd J. 2014. Commingled Human Remains. Methods in Recovery, Analysis, and Identification. First edition. Oxford: Academic Press.

Adserias-Garriga J. 2019. Age Estimation: A Multidisciplinary Approach. First edition. London: Academic Press.

Adserias-Garriga J, Wilson-Taylor R. 2019. Skeletal age estimation in adults. In: Adserias-Garriga J, editor. Age Estimation: A Multidisciplinary Approach. First edition. London: Academic Press. p 55–73.

Adserias-Garriga J, Zapico S. 2018. Age assessment in forensic cases: Anthropological, odontological and biochemical methods for age estimation in the dead. M J Foren 1:1.

Algee-Hewitt BFB. 2017. Age Estimation in Modern Forensic Anthropology. In: Langley N, Tersigni-Tarrant MA, editors. Forensic Anthropology: A Comprehensive Introduction. Second edition. Boca Raton: CRC Press. p 381–419.

Alicioglu B, Kartal O, Gurbuz H, Sut N. 2008. Symphysis pubis distance in adults: A retrospective computed tomography study. Surg Radiol Anat 30:153–157.

AlQahtani S. 2019. Dental age estimation in fetal and children. In: Adserias-Garriga, editor. Age Estimation: A Multidisciplinary Approach. First edition. London: Academic Press. p 89–106.

AlQahtani SJ, Hector MP, Liversidge HM. 2010. Brief communication: The London atlas of human tooth development and eruption. Am J Phys Anthropol 142:481–490.

Alves-Cardoso F, Assis S. 2018. Can osteophytes be used as age at death estimators? Testing correlations in skeletonized human remains with known age-at-death. Forensic Sci Int 288:59–66.

Anderson MF, Anderson DT, Wescott DJ. 2010. Estimation of adult skeletal age-at-death using the sugeno fuzzy integral. Am J Phys Anthropol 142:30–41.

Arany S, Ohtani S. 2010. Age estimation by racemization method in teeth: Application of aspartic acid, glutamate, and alanine. J Forensic Sci 55:701–705.

Arden N, Nevitt MC. 2006. Osteoarthritis: Epidemiology. Best Pract Res Clin Rheumatol 20:3–25.

Aykroyd RG, Lucy D, Pollard AM, Roberts CA. 1999. Nasty, brutish, but not necessarily short: a reconsideration of the statistical methods used to calculate age at death from adult human skeletal and dental age indicators. Am Antiq 64:55–70.

Baccino E, Cunha E, Cattaneo C. 2013. Aging the Dead and the Living. In: Siegel J, Saukko P, editors. Encyclopedia of Forensic Sciences. Waltham: Academic Press. p 42–48.

Baccino E, Schmitt A. 2006. Determination of adult age at death in the forensic context. In: Schmitt A, Cunha E, Pinheiro J, editors. Forensic anthropology and medicine: complementary sciences from recovery to cause of death. Totowa: Humana Press Inc. p 259–280.

Baccino E, Sinfield L, Colomb S, Baum TP, Martrille L. 2014. Technical note: The two step procedure (TSP) for the determination of age at death of adult human remains in forensic cases. Forensic Sci Int 244:247–251.

Baccino E, Tavernier JC, Lamendin H, Frammery D, Nossintchouk R, Humbert JF. 1991. Recherche d'une méthode multifactorielle simple pour la détermination de l'age des cadavres adultes. J Med Leg Droit 34:27–33.

Baccino E, Ubelaker D, Hayek L, Zerilli A. 1999. Evaluation of seven methods of estimating age at death from mature human skeletal remains. J Forensic Sci 44:931–936.

Baccino E, Zerilli A. 1997. The two step strategy (TSS) or the right way to combine a dental (Lamendin) and an anthropological (Suchey–Brooks system) method for age determination. Am Acad Forensic Sci. New York. p 150.

Barrier P, Dedouit F, Braga J, Joffre F, Rougé D, Rousseau H, Telmon N. 2009. Age at death estimation using multislice computed tomography reconstructions of the posterior pelvis. J Forensic Sci 54:773–778.

Bartolini V, Pinchi V, Gualco B, Vanin S, Chiaracane G, D'Elia G, Norelli GA, Focardi M. 2018. The iliac crest in forensic age estimation: evaluation of three methods in pelvis X-rays. Int J Legal Med 132:279–288.

Bassed RB, Drummer OH, Briggs C, Valenzuela A. 2011. Age estimation and the medial clavicular epiphysis: Analysis of the age of majority in an Australian population using computed tomography. Forensic Sci Med Pathol 7:148–154.

Becker I, Woodley SJ, Stringer MD. 2010. The adult human pubic symphysis: A systematic review. J Anat 217:475–487.

Bedford ME, Russell KF, Lovejoy CO, Meindl RS, Simpson SW, Stuart-Macadam PL. 1993. Test of the multifactorial aging method using skeletons with known ages-at-death from the grant collection. Am J Phys Anthropol 91:287–297.

Bekaert B, Kamalandua A, Zapico SC, Van De Voorde W, Decorte R. 2015. Improved age determination of blood and teeth samples using a selected set of DNA methylation markers. Epigenetics 10:922–930.

Belcastro G, Rastelli E, Mariotti V. 2008. Variation of the Degree of Sacral Vertebral Body Fusion in Adulthood in Two European Modern Skeletal Collections. Am J Phys Anthropol 135:149–160.

Berenbaum F, Sellam J. 2008. Obesity and osteoarthritis: what are the links? Joint Bone Spine 75:667–668.

Berg G. 2017. Sex Estimation of Unknown Human Skeletal Remains. In: Langley N, Tersigni-Tarrant MA, editors. Forensic Anthropology: A Comprehensive Introduction. Second edition. Boca Raton: CRC Press. p 143–161.

Berg GE. 2008. Pubic bone age estimation in adult women. J Forensic Sci 53:569–577.

Bertrand B, Cunha E, Bécart A, Gosset D, Hédouin V. 2019a. Age at death estimation by cementochronology: Too precise to be true or too precise to be accurate? Am J Phys Anthropol 169:464–481.

Bertrand B, Oliveira-Santos I, Cunha E. 2019b. Cementochronology: a validated but disregarded method for age at death estimation. In: Adserias-Garriga J, editor. Age Estimation: A Multidisciplinary Approach. First edition. London: Academic Press. p 169–186.

Bethard JD. 2005. A Test of the Transition Analysis Method for Estimation of Age-at-Death in Adult Human Skeletal Remains. Master thesis. University of Tennessee.

Bilfeld MF, Dedouit F, Sans N, Rousseau H, Rougé D, Telmon N. 2013. Ontogeny of size and shape sexual dimorphism in the ilium: A multislice computed tomography study by geometric morphometry. J Forensic Sci 58:303–310.

Biwasaka H, Aoki Y, Takahashi Y, Fukuta M, Usui A. 2019. A quantitative morphological analysis of three-dimensional CT coxal bone images of contemporary Japanese using homologous models for sex and age estimation. Leg Med 36:1–8.

Biwasaka H, Sato K, Aoki Y, Kato H, Maeno Y, Tanijiri T, Fujita S, Dewa K. 2013. Three dimensional surface analyses of pubic symphyseal faces of contemporary Japanese reconstructed with 3D digitized scanner. Leg Med 15:264–268.

Blenkin MRB, Evans W. 2010. Age estimation from the teeth using a modified demirjian system. J Forensic Sci 55:1504–1508.

Bocklandt S, Lin W, Sehl M, Sánchez F, Sinsheimer J, Horvath S, Vilain E. 2011. Epigenetic predictor of age. PLoS One 6:e14821.

Bocquet-Appel JP, Masset C. 1982. Farewell to paleodemography. J Hum Evol 11:321–333.

Boldsen JL, Milner GR, Konigsberg LW, Wood J. 2002. Transition analysis: a new method for estimating age from skeletons. In: Hoppa R, Vaupel J, editors. Paleodemography: Age distributions from skeletal samples. First edition. Cambridge: Cambridge University Press. p 73–106.

Botha D, Lynnerup N, Steyn M. 2019. Age estimation using bone mineral density in South Africans. Forensic Sci Int 297:307–314.

Brennaman AL, Love KR, Bethard JD, Pokines JT. 2017. A bayesian approach to age-at-death estimation from osteoarthritis of the shoulder in modern North Americans. J Forensic Sci

62:573–584.

Brooks S, Suchey JM. 1990. Skeletal age determination based on the os pubis: a comparison of the Acsádi- Nemeskéri and Suchey-Brooks methods. Hum Evol 5:227–238.

Browning RC, Kram R. 2007. Effects of obesity on the biomechanics of walking at different speeds. Med Sci Sports Exerc 39:1632–1641.

Brůžek J. 1995. Diagnose sexuelle à l'aide de l'analyse discriminante appliquée au tibia. Antropol Port 13:93–106.

Brůžek J. 2002. A method for visual determination of sex, using the human hip bone. Am J Phys Anthropol 117:157–168.

Brůžek J, Murail P. 2006. Methodology and reliability of sex determination from the skeleton. In: Schmitt A, Cunha E, Pinheiro J, editors. Forensic Anthropology and Medicine: Complementary Sciences From Recovery to Cause of Death. First edition. Totowa: Humana Press Inc. p 225–242.

Brůžek J, Santos F, Dutailly B, Murail P, Cunha E. 2017. Validation and reliability of the sex estimation of the human os coxae using freely available DSP2 software for bioarchaeology and forensic anthropology. Am J Phys Anthropol 164:440–449.

Buckberry JL, Chamberlain AT. 2002. Age Estimation From the Auricular Surface of the Ilium: A Revised Method. Am J Phys Anthropol 119:231–239.

Buk Z, Kordík P, Brůžek J, Schmitt A, Šnorek M. 2012. The age at death assessment in a multi-ethnic sample of pelvic bones using nature-inspired data mining methods. Forensic Sci Int 220:294.e1-e9.

Bullock M, Márquez L, Hernández P, Ruíz F. 2013. Paleodemographic age-at-death distributions of two Mexican skeletal collections: A comparison of transition analysis and traditional aging methods. Am J Phys Anthropol 152:67–78.

Le Cabec A, Tang NK, Rubio V, Hillson S. 2019. Nondestructive adult age at death estimation : Visualizing cementum annulations in a known age historical human assemblage using synchrotron X-ray microtomography. Am J Phys Anthropol 168:25–44.

Cabo LL, Brewster C, Azpiazu J. 2012. Sexual Dimorphism: Interpreting Sex Markers. In: Dirkmaat DC, editor. A Companion to Forensic Anthropology. First edition. London: Wiley-Blackwell. p 248–286.

Calce SE. 2012. A new method to estimate adult age-at-death using the acetabulum. Am J Phys Anthropol 148:11–23.

Calce SE, Kurki HK, Weston DA, Gould L. 2018a. The relationship of age, activity, and body size on osteoarthritis in weight-bearing skeletal regions. Int J Paleopathol 22:45–53.

Calce SE, Kurki HK, Weston DA, Gould L. 2018b. The effects of osteoarthritis on age-at-death estimates from the human pelvis. Am J Phys Anthropol 167:3–19.

Calce SE, Rogers TL. 2011. Evaluation of age estimation technique: Testing traits of the acetabulum to estimate age at death in adult males. J Forensic Sci 56:302–311.

Campanacho V, Santos AL, Cardoso HFV. 2012. Assessing the influence of occupational and physical activity on the rate of degenerative change of the pubic symphysis in Portuguese males from the 19th to 20th century. Am J Phys Anthropol 148:371–378.

Campanacho VC. 2016. The influence of skeletal size on age-related criteria from the pelvic joints in Portuguese and North American samples. Dissertation thesis. University of Sheffield.

Cantín M, Muñoz M, Olate S. 2015. Generation of 3D tooth models based on three-dimensional scanning to study the morphology of permanent teeth. Int J Morphol 33:782–787.

Cappella A, Cummaudo M, Arrigoni E, Collini F, Cattaneo C. 2017. The issue of age estimation in a modern skeletal population: are even the more modern current aging methods satisfactory for the elderly? J Forensic Sci:12–17.

Cardoso HF V, Ríos L. 2011. Age estimation from stages of epiphyseal union in the presacral vertebrae. Am J Phys Anthropol 144:238–247.

Caruso S, Bernardi S, Pasini M, Giuca MR, Docimo R, Continenza MA, Gatto R. 2016. The process of mineralisation in the development of human tooth. Eur J Paediatr Dent 17:322–326.

Cattaneo C. 2007. Forensic anthropology: developments of a classical discipline in the new millennium. Forensic Sci Int 165:185–93.

Cavaignac E, Li K, Faruch M, Savall F, Chiron P, Huang W, Telmon N. 2017. Three-dimensional geometric morphometric analysis reveals ethnic dimorphism in the shape of the femur. J Exp othopaedics 4:13.

Čechová M, Dupej J, Brůžek J, Bejdová Š, Horák M, Velemínská J. 2019. Sex estimation using external morphology of the frontal bone and frontal sinuses in a contemporary Czech population. Int J Legal Med 133:1285–1294.

Chaillet N, Willems G. 2004. Dental maturity in Belgian children using Demirjian's method and polynomial functions: New standard curves for forensic and clinical use. J Forensic Odontostomatol 22:18–27.

Chandna S, Bathla M. 2011. Oral manifestations of thyroid disorders and its management. Indian J Endocrinol Metab 15:113–116.

Chapman T, Lefevre P, Semal P, Moiseev F, Sholukha V, Louryan S, Rooze M, Sint S Van. 2014. Sex determination using the Probabilistic Sex Diagnosis (DSP : Diagnose Sexuelle Probabiliste) tool in a virtual environment. Forensic Sci Int 234:189-e1.

Chiba F, Makino Y, Motomura A, Inokuchi G, Torimitsu S, Ishii N, Kubo Y, Abe H, Sakuma A, Nagasawa S, Saitoh H, Yajima D, Hayakawa M, Miura M, Iwase H. 2014. Age estimation by quantitative features of pubic symphysis using multidetector computed tomography. Int J

Legal Med 128:667–673.

Christensen AM, Passalacqua N V., Bartelink EJ. 2014a. Ancestry Estimation. In: Christensen AM, Passalacqua N V, Bartelink EJ, editors. Forensic Anthropology: Current Methods and Practice. First edition. San Diego: Academic Press. p 223–242.

Christensen AM, Passalacqua N V., Bartelink EJ. 2014b. Age estimation. In: Christensen AM, Passalacqua N V, Bartelink EJ, editors. Forensic Anthropology: Current Methods and Practice. First edition. San Diego: Academic Press. p 243–284.

Christensen AM, Passalacqua N V., Bartelink EJ. 2019. Age estimation: Current methods and practice. In: Christensen AM, Passalacqua N V, Bartelink EJ, editors. Forensic Anthropology. Second edition. San Diego: Academic Press. p 307–349.

Christensen AM, Passalacqua N V, Bartelink EJ. 2014c. Forensic Anthropology: Current Methods and Practice. First edition. San Diego: Academic Press.

Cignoni P, Callieri M, Corsini M, Dellepiane M, Ganovelli F, Ranzuglia G. 2008. MeshLab: an Open-Source Mesh Processing Tool. In 6th Eurographics Ital Chapter Conf, pp 129–136.

Čihák R. 2011. Anatomie I. Third edition. Grada publishing.

Citardi MJ, Herrmann B, Hollenbeak CS, Stack BC, Cooper M, Bucholz RD. 2001. Comparison of scientific calipers and computed-enabled CT review for the measurement of skull base and craniomaxillofacial dimensions. Skull base 11:5–11.

Colard T, Bertrand B, Naji S, Delannoy Y, Bécart A. 2015. Toward the adoption of cementochronology in forensic context. Int J Legal Med 132:1117–1124.

Colman KL, Merwe AE Van Der, Stull KE, Dobbe JGG, Streekstra GJ, Rijn RR Van, Oostra R, Boer HH De, Colman KL. 2019. The accuracy of 3D virtual bone models of the pelvis for morphological sex estimation. Int J Legal Med 133:1853–1860.

Corron L, Marchal F, Condemi S, Chaumoître K, Adalian P. 2017. Evaluating the consistency, repeatability, and reproducibility of osteometric data on dry bone surfaces, scanned dry bone surfaces, and scanned bone surfaces obtained from living individuals. Bull Mem Soc Anthropol Paris 29:33–53.

Couoh LR. 2017. Differences between biological and chronological age-at-death in human skeletal remains: A change of perspective. Am J Phys Anthropol 163:671–695.

Crowder C, Andronowski J, Dominguez V. 2018. Bone histology as an integrated tool in the process of human identification. In: Latham KE, Bartelink EJ, Finnegan M, editors. New perspectives in forensic human skeletal identification. First edition. London: Academic Press. p 201–213.

Crowder C, Heinrich J, Stout S. 2012. Rib histomorphometry for adult age estimation. In: Lynne S, Bell L, editors. Forensic microscopy for skeletal tissues: methods and protocols. Methods in molecular biology. First edition. London: Humana Press. p 109–127.

Cunha E, Baccino E, Martrille L, Ramsthaler F, Prieto J, Schuliar Y, Lynnerup N, Cattaneo C. 2009. The problem of aging human remains and living individuals: a review. Forensic Sci Int 193:1–13.

Cunha E, Wasterlain S. 2007. The Coimbra identified osteological collections. In: Grupe G, Peters J, editors. Skeletal Series in their Socioeconomic Context. Documenta Archaeobiologiae. Vol. 5. Rahden: Verlag Marie Leidorf. p 23–33.

Cunningham CA. 2019. Skeletal age estimation in fetal, infant, children and subadults. In: Adserias-Garriga J, editor. Age Estimation: A Multidisciplinary Approach. First edition. London: Academic Press. p 41–54.

Cunningham CA, Scheuer L, Black S. 2016. Developmental Juvenile Osteology. Second edition. London: Academic Press.

Daubert v. Merrell Dow Pharmaceuticals, Inc. 509 US 579, 1993.

Dayal MR, Kegley ADT, Štrkalj G, Bidmos MA, Kuykendall KL. 2009. The history and composition of the Raymond A. Dart collection of human skeletons at the University of the Witwatersrand, Johannesburg, South Africa. Am J Phys Anthropol 140:324–335.

Decker SJ, Davy-Jow SL, Ford JM, Hilbelink DR. 2011. Virtual determination of sex: metric and nonmetric traits of the adult pelvis from 3D computed tomography models. J Forensic Sci 56:1107–14.

Dedouit F, Telmon N, Costagliola R, Otal P, Joffre F, Rougé D. 2007. Virtual anthropology and forensic identification: Report of one case. Forensic Sci Int 173:182–187.

Demirjian A, Goldstein H. 1976. New systems for dental maturity based on seven and four teeth. Ann Hum Biol 3:411–421.

Demirjian A, Goldstein H, Tanner JM. 1973. A new system of dental age assessment. Hum Biol 45:211–227.

DeWitte S. 2017. Demographic anthropology. Am J Phys Anthropol 165:893–903.

Dias PEM, Beaini TL, Melani RFH. 2010. Age estimation from dental cementum incremental lines and periodontal disease. J Forensic Odontostomatol 28:13–21.

Digangi EA, Bethard JD, Kimmerle EH, Konigsberg LW. 2009. A new method for estimating age-at-death from the first rib. Am J Phys Anthropol 138:164–176.

Dirkmaat D. 2012. A companion to forensic anthropology. First edition. London: Wiley-Blackwell.

Donato L, Cipolloni L, Ozonoff A, di Luca A. 2016. A preliminary study of the relationship between obliteration of cranial sutures and age at time of death. Biol Syst Open Access 5:36–38.

Draft DM, Kasper KA. 2019. Dental age assessment in late adolescence. In: Adserias-Garriga J, editor. Age Estimation: A Multidisciplinary Approach. First edition. London: Academic Press. p 107–123.

Dubourg O, Faruch-Bilfeld M, Telmon N, Maupoint E, Saint-Martin P, Savall F. 2019. Correlation between pubic bone mineral density and age from a computed tomography sample. Forensic Sci Int 298:345–350.

Dudar JC. 1993. Identification of rib number and assessment of intercostal variation at the sternal rib end. J Forensic Sci 38:788–797.

Dudzik B, Langley NR. 2015. Estimating age from the pubic symphysis: A new component-based system. Forensic Sci Int 257:98–105.

Ekizoglu O, Hocaoglu E, Inci E, Sayin I, Solmaz D, Bilgili MG, Can IO. 2014. Forensic age estimation by the Schmeling method: computed tomography analysis of the medial clavicular epiphysis. Int J Legal Med 129:203–210.

Ellingham S, Adserias-Garriga J. 2019. Complexities and considerations of human age estimation. In: Adserias-Garriga J, editor. Age Estimation: A Multidisciplinary Approach. First edition. London: Academic Press. p 1–16.

Errickson D. 2017. Shedding light on skeletal remains: The use of structured light scanning for 3D archiving. In: Errickson D, Thompson T, editors. Human remains: Another dimension. First edition. Academic Press. p 93–101.

Falys CG, Lewis ME. 2011. Proposing a way forward: A review of standardisation in the use of age categories and ageing techniques in osteological analysis (2004-2009). Int J Osteoarchaeol 21:704–716.

Falys CG, Prangle D. 2015. Estimating age of mature adults from the degeneration of the sternal end of the clavicle. Am J Phys Anthropol 156:203–214.

Falys CG, Schutkowski H, Weston DA. 2006. Auricular surface aging: Worse than expected? A test of the revised method on a documented historic skeletal assemblage. Am J Phys Anthropol 130:508–513.

Fazekas IG, Kósa F. 1978. Forensic fetal osteology. Budapest: Akadémiai Kiadó.

Ferembach D, Schwidetzky I, Stloukal M. 1980. Recommendations for age and sex diagnoses of skeletons. J Hum Evol 9:517–549.

Ferrant O, Rougé-Maillart C, Guittet L, Papin F, Clin B, Fau G, Telmon N. 2009. Age at death estimation of adult males using coxal bone and CT scan: A preliminary study. Forensic Sci Int 186:14–21.

Ferreira MT, Vicente R, Navega D, Gonçalves D, Curate F, Cunha E. 2014. A new forensic collection housed at the University of Coimbra, Portugal: The 21st century identified skeletal collection. Forensic Sci Int 245:202.e1–202.e5.

Fojas CL, Kim J, Minsky-Rowland JD, Algee-Hewitt BFB. 2018. Testing inter-observer reliability of the Transition Analysis aging method on the William M. Bass forensic skeletal collection. Am J Phys Anthropol 165:183–193.

Franklin D. 2010. Forensic age estimation in human skeletal remains: Current concepts and future directions. Leg Med 12:1–7.

Franklin D, Cardini A, Flavel A, Kuliukas A. 2013. Estimation of sex from cranial measurements in a Western Australian population. Forensic Sci Int 229:158.e1-158.e8.

Franklin D, Flavel A. 2014. Brief Communication : Timing of Spheno-Occipital Closure in Modern Western Australians. Am J Phys Anthropol 153:132–138.

Franklin D, Flavel A. 2019. Population specificity in the estimation of skeletal age and sex : case studies using a Western Australian population. Aust J Forensic Sci 51:S188-192.

Franklin D, Flavel A, Noble J, Swift L, Karkhanis S. 2015. Forensic age estimation in living individuals: methodological considerations in the context of medico-legal practice. Res Reports Forensic Med Sci 5:53–66.

Friedlaender JS, Costa PT, Bosse R, Ellis E, Rhoads JG, Stoudt HW. 1977. Longitudinal physique changes among healthy white veterans at Boston. Hum Biol 49:541–558.

Friess M. 2012. Scratching the surface? The use of surface scanning in physical and paleoanthropology. J Anthropol Sci 90:1–25.

Fully G. 1956. Une nouvelle methode de determination de la taille. Ann med Leg 35:266–273.

Galasso O, Familiari F, De Gori M, Gasparini G. 2012. Recent findings on the role of gelatinases (Matrix Metalloproteinase-2 and -9) in osteoarthritis. Adv Orthop 2012:1–7.

Galera V, Ubelaker DH, Hayek L-AC. 1998. Comparison of macroscopic cranial methods of age estimation applied to skeletons from the Terry Collection. J Forensic Sci 43:933–939.

Garamendi PM, Landa MI, Botella MC, Alemán I. 2011. Forensic age estimation on digital x-ray images: Medial epiphyses of the clavicle and first rib ossification in relation to chronological age. J Forensic Sci 56:3–12.

Garvin HM. 2012. Adult Sex Determination: Methods and Application. In: Dirkmaat DC, editor. A Companion to Forensic Anthropology. First edition. London: Wiley-Blackwell. p 239–247.

Garvin HM, Passalacqua N V. 2012. Current practices by forensic anthropologists in adult skeletal age estimation. J Forensic Sci 57:427–433.

Garvin HM, Uhl N, Passalacqua N V., Gipson D, Overbury RS, Cabo L. 2012. Developments in forensic anthropology: Age at- death estimation. in: Dirkmaat DC, editor. A companion to forensic anthropology. First edition. London: Wiley-Blackwell. p 202–223.

Gatta EA, Al-alousi WS, Diab BS. 2008. Primary teeth emergence in relation to nutritional status among 4-48 months old children in Baghdad city. Mustansiria Dent J 5:62–70.

Gaur R, Kumar P. 2012. Effect of undernutrition on deciduous tooth emergence among Rajput children of Shimla District of Himachal Pradesh, India. Am J Phys Anthropol 148:54–61.

Gilbert BM, McKern TW. 1973. A method for aging the female os pubis. Am J Phys Anthropol 38:31–38.

Giuliani C, Cilli E, Bacalini MG, Pirazzini C, Sazzini M, Gruppioni G, Franceschi C, Garagnani P, Luiselli D. 2016. Inferring chronological age from DNA methylation patterns of human teeth. Am J Phys Anthropol 159:585–595.

Gocha TP, Ingvoldstad ME, Kolatorowicz A, Cosgriff-Hernandez MTJ, Sciulli PW. 2015. Testing the applicability of six macroscopic skeletal aging techniques on a modern Southeast Asian sample. Forensic Sci Int 249:318-e1.

Godde K, Hens SM. 2012. Age-at-death estimation in an Italian historical sample: A test of the Suchey-Brooks and transition analysis methods. Am J Phys Anthropol 149:259–265.

Godde K, Hens SM. 2015. Modeling senescence changes of the pubic symphysis in historic italian populations: A comparison of the rostock and forensic approaches to aging using transition analysis. Am J Phys Anthropol 156:466–473.

Goliath JR, Stewart MC, Stout SD. 2016. Variation in osteon histomorphometrics and their impact on age-at-death estimation in older individuals. Forensic Sci Int 262:282.e1-282.e6.

Grabherr S, Cooper C, Ulrich-Bochsler S, Uldin T, Ross S, Oesterhelweg L, Bolliger S, Christe A, Schnyder P, Mangin P, Thali MJ. 2009. Estimation of sex and age of “virtual skeletons”-a feasibility study. Eur Radiol 19:419–429.

Grivas CR, Komar DA. 2008. Kumho, Daubert, and the nature of scientific inquiry: Implications for forensic anthropology. J Forensic Sci 53:771–776.

Grosskopf B, McGlynn G. 2011. Age diagnosis based on incremental lines in dental cementum: A critical reflection. Anthropol Anzeiger 68:275–289.

Gurses MS, Inanir NT, Gokalp G, Fedakar R, Tobcu E, Ocakoglu G. 2016. Evaluation of age estimation in forensic medicine by examination of medial clavicular ossification from thin-slice computed tomography images. Int J Legal Med 130:1343–1352.

Gustafson G. 1950. Age Determinations on Teeth. J Am Dent Assoc 41:45–54.

Guydish M, Henson K. 2017. Using digitized Native American skeletal remains to conduct osteological analyses. Proc W Va Acad Sci, 89.

Hagelthorn CL, Alblas A, Greyling L. 2019. The accuracy of the Transition Analysis of aging on a heterogenic South African population. Forensic Sci Int 297:370.e1-370.e5.

Hanihara K, Suzuki T. 1978. Estimation of age from the pubic symphysis by means of multiple regression analysis. Am J Phys Anthropol 48:233–240.

Hartnett K, Fulginiti L, Seidel A. 2018. Adult age-at-death estimation in unknown decedents: New perspectives on an old problem. In: Latham K, Bartelink E, Finnegan M, editors. New perspectives in forensic human skeletal identification. First edition. London: Academic Press. p 65–85.

Hartnett KM. 2010a. Analysis of age-at-death estimation using data from a new, modern autopsy sample - Part II: Sternal end of the fourth rib. J Forensic Sci 55:1152–1156.

Hartnett KM. 2010b. Analysis of age-at-death estimation using data from a new, modern autopsy sample - Part I: Pubic bone. J Forensic Sci 55:1145–1151.

Hefner JT. 2009. Cranial nonmetric variation and estimating ancestry. J Forensic Sci 54:985–995.

Hefner JT, Ousley SD. 2014. Statistical classification methods for estimating ancestry using morphoscopic traits. J Forensic Sci 59:883–890.

Hefner JT, Spradley MK, Anderson B. 2014. Ancestry assessment using random forest modeling. J Forensic Sci 59:583–589.

Hillewig E, Degroote J, Van Der Paelt T, Visscher A, Vandemaele P, Lutin B, D'Hooghe L, Vandriessche V, Piette M, Verstraete K. 2013. Magnetic resonance imaging of the sternal extremity of the clavicle in forensic age estimation: Towards more sound age estimates. Int J Legal Med 127:677–689.

Hisham S, Abdullah N, Noor M, Franklin D. 2019. Quantification of pubic symphysis metamorphosis based on the analysis of clinical MDCT scans in a contemporary Malaysian population. J Forensic Sci 64:1803–1811.

Hisham S, Flavel A, Abdullah N, Franklin D. 2018. Quantification of spheno-occipital synchondrosis fusion in a contemporary male Malaysian population. Forensic Sci Int 284:78–84.

Hoffman J. 1979. Age estimations from diaphyseal lengths: two months to twelve year. J Forensic Sci 24:461–469.

Hoppa R, Vaupel J. 2002. The Rostock Manifesto for paleodemography: the way from stage to age. In: Hoppa R, Vaupel J, editors. Paleodemography: age distributions from skeletal samples. First edition. Cambridge: Cambridge University Press. p 1–8.

Hoppa RD. 2000. Population variation in osteological aging criteria: An example from the pubic symphysis. Am J Phys Anthropol 111:185–191.

Hora M, Sládek V. 2018. Population specificity of sex estimation from vertebrae. Forensic Sci Int 291:279.e1-279.e12.

İşcan M, Steyn M. 2013a. The human skeleton in forensic medicine. Third edition. Springfield: Charles C Thomas.

İşcan MY, Loth SR, Wright RK. 1984. Age Estimation from the Rib by Phase Analysis: White Males. J Forensic Sci 29:1094–1104.

İşcan MY, Loth SR, Wright RK. 1985. Age Estimation from the Rib by Phase Analysis: White Females. J Forensic Sci 30:853–863.

İşcan MY, Miller-Shaivitz P. 1984. Discriminant function sexing of the tibia. J Forensic Sci 29:1087–1093.

İşcan YM, Loth SR, Wright RK. 1984b. Metamorphosis at the Sternal Rib End: A New Method to Estimate Age at Death in White Males. Am J Phys Anthropol 65:147–156.

İşcan YM, Steyn M. 2013a. Sex. In: The Human Skeleton in Forensic Medicine. Third edition. Springfield: Charles C Thomas. p 143–193.

İşscan YM, Steyn M. 2013b. Ancestry. In: The Human Skeleton in Forensic Medicine. Third edition. Springfield: Charles C Thomas. p 195–226.

İşcan YM, Steyn M. 2013c. Skeletal age. In: The Human Skeleton in Forensic Medicine. Third edition. Springfield: Charles C Thomas. p 59–141.

İşcan YM, Steyn M. 2013d. Dental analysis. In: The Human Skeleton in Forensic Medicine. Third edition. Springfield: Charles C Thomas. p 259–289.

İşcan YM, Steyn M. 2013b. Stature. In: The Human Skeleton in Forensic Medicine. Third edition. Springfield: Charles C Thomas. p 227–258.

Jafarzadeh SR, Felson DT. 2018. Updated estimates suggest a much higher prevalence of arthritis in United States adults than previous ones. Arthritis Rheumatol 70:185–192.

Jantz Meadows L, Jantz RL. 1999. Secular change in long bone length and proportion in the United States, 1800 – 1970. Am J Phys Anthropol 110:57–67.

Johnson L, Bethard J. 2019. Testing a computational approach for estimating age-at-death on a modern Portuguese population. In: 88th Annual meeting of the American association of physical anthropologists. Cleveland, p 116.

Jooste N, L'Abbé EN, Pretorius S, Steyn M. 2016. Validation of transition analysis as a method of adult age estimation in a modern South African sample. Forensic Sci Int 266:580.e1-580.e7.

Joubert LC, Briers N, Meyer A. 2019. Evaluation of the enhanced computational methods of estimating age-at-death using the pubic symphyses of a white South African population. J Forensic Sci:1–9.

Kasetty S, Rammanohar M, Raju Ragavendra T. 2010. Dental cementum in age estimation: A polarized light and stereomicroscopic study. J Forensic Sci 55:779–783.

Katherine Spradley M, Jantz RL. 2016. Ancestry estimation in forensic anthropology: Geometric morphometric versus standard and nonstandard interlandmark distances. J Forensic Sci 61:892–897.

Katz D, Suchey JM. 1986. Age determination of the male os pubis. Am J Phys Anthropol 69:427–435.

Kellinghaus M, Schulz R, Vieth V, Schmidt S, Schmeling A. 2010. Forensic age estimation in living subjects based on the ossification status of the medial clavicular epiphysis as revealed by thin-slice multidetector computed tomography. Int J Legal Med 124:149–154.

Kim J. 2016. Understanding population-specific age estimation using documented Asian skeletal samples. In: 86th Annual meeting of the American association of physical anthropologists. New Orleans, p 243.

Kim J, Algee-Hewitt BFB, Stoyanova DK, Figueroa-Soto C, Slice D. 2018. Testing reliability of

the computational age-at-death estimation methods between five observers using three-dimensional image data of the pubic symphysis. J Forensic Sci 64:507–518.

Kimmerle EH, Konigsberg LW, Jantz RL, Baraybar JP. 2008. Analysis of age-at-death estimation through the use of pubic symphyseal data. J Forensic Sci 53:558–568.

Klein S, Avery M, Adams G, Pollard S, Simske S. 2014. From scan to print: 3D printing as a means for replication. In: NIP & Digital fabrication conference. Society for imagining science and technology. p 417–421.

Konigsberg LW, Algee-Hewitt BFB, Steadman DW. 2009. Estimation and evidence in forensic anthropology: sex and race. Am J Phys Anthropol 139:77–90.

Konigsberg LW, Frankenberg SR. 1992. Estimation of age structure in anthropological demography. Am J Phys Anthropol 89:235–256.

Kotěrová A, Králík V, Rmoutilová R, Friedl L, Růžička P, Velemínská J, Marchal F, Brůžek J. 2019. Impact of 3D surface scanning protocols on the os coxae digital data: implications for sex and age-at-death assessment. J Forensic Leg Med 68: 101866.

Kotěrová A, Navega D, Štepanovský M, Buk Z, Brůžek J, Cunha E. 2018a. Age estimation of adult human remains from hip bones using advanced methods. Forensic Sci Int 287:163–175.

Kotěrová A, Velemínská J, Cunha E, Brůžek J. 2018b. A validation study of the Stoyanova et al. method (2017) for age-at-death estimation quantifying the 3D pubic symphyseal surface of adult males of European populations. Int J Legal Med 133:603–612.

Kotěrová A, Velemínská J, Dupej J, Brzobohatá H, Pilný A, Brůžek J. 2017. Disregarding population specificity: its influence on the sex assessment methods from the tibia. Int J Legal Med 131:251–261.

Van der Kraan PM, van den Berg WB. 2007. Osteophytes: relevance and biology. Osteoarthr Cartil 15:237–244.

Kranioti EF, Apostol M a. 2014. Sexual dimorphism of the tibia in contemporary Greeks, Italians, and Spanish: forensic implications. Int J Legal Med 129:357–363.

Kranioti EF, İşcan MY, Michalodimitrakis M. 2008. Craniometric analysis of the modern Cretan population. Forensic Sci Int 180: 110.e111–110.e115.

Kranioti EF, Michalodimitrakis M. 2009. Sexual dimorphism of the humerus in contemporary Cretans - A population-specific study and a review of the literature. J Forensic Sci 54:996–1000.

Kreitner KF, Schweden FJ, Riepert T, Nafe B, Thelen M. 1998. Bone age determination based on the study of the medial extremity of the clavicle. Eur Radiol 8:1116–1122.

Kunos CA, Simpson SW, Russell KF, Hershkovitz I. 1999. First rib metamorphosis: Its possible utility for human age-at-death estimation. Am J Phys Anthropol 110:303–323.

Kuzminsky SC, Gardiner MS. 2012. Three-dimensional laser scanning : potential uses for museum

conservation and scienti fi c research. J Archaeol Sci 39:2744–2751.

L'Abbé EN, Kenyhercz M, Stull KE, Keough N, Nawrocki S. 2013. Application of Fordisc 3.0 to explore differences among crania of North American and South African blacks and whites. J Forensic Sci 58:1579–1583.

L´Abbé E, Steyn M. 2012. The establishment and advancement of forensic anthropology South Africa. In: Dirkmaat DC, editor. A Companion to Forensic Anthropology. First edition. London: Wiley-Blackwell. p 626–638.

Lagacé F, Verna E, Adalian P, Baccino E, Martrille L. 2019. Testing the accuracy of a new histomorphometric method for age-at-death estimation. Forensic Sci Int 296:48–52.

Lamendin H. 1988. Appréciation d'age par la méthode de Gustafson <<simplifiée>>. Chir Dent Fr 427:205–214.

Lamendin H, Baccino E, Humbert JF, Tavernier JC, Nossintchouk RM, Zerilli A. 1992. A simple technique for age estimation in adult corpses: The two criteria dental method. J Forensic Sci 37:1373–1379.

Langley-Shirley N, Jantz RL. 2010. A bayesian approach to age estimation in modern Americans from the clavicle. J Forensic Sci 55:571–583.

Langley NR. 2017. Stature estimation. In: Langley NR, Tersigni-Tarrant MA, editors. Forensic Anthropology: A Comprehensive Introduction. Second edition. Boca Raton: CRC Press. p 195–203.

Langley NR, Dudzik B, Cloutier A. 2018. A decision tree for nonmetric sex assessment from the skull. J Forensic Sci 63:31–37.

Langley NR, Gooding AF, Tersigni-Tarrant MA. 2017. Age Estimation Methods. In: Langley NR, Tersigni-Tarrant MA, editors. Forensic Anthropology: A Comprehensive Introduction. Second edition. Boca Raton: CRC Press. p 175–194.

Langley NR, Tersigni-Tarrant MA. 2017. Forensic Anthropology: A Comprehensive Introduction. Second edition. Boca Raton: CRC Press.

Lanteri L, Bizot B, Saliba-Serre B, Gaudart J, Signoli M, Schmitt A. 2018. Cementochronology: A solution to assess mortality profiles from individual age-at-death estimates. J Archaeol Sci Reports 20:576–587.

Latham K, Bartelink E, Finnegan M. 2018. New perspectives in forensic human skeletal identification. First edition. London: Academic Press.

Lei G, Liu F, Liu P, Zhou Y, Jiao T, Dang Y. 2019. Worldwide tendency and focused research in forensic anthropology : A bibliometric analysis of decade (2008–2017). Leg Med 37:67–75.

Lewis J, Kasper K. 2018. Assessment of Dental Age. In: Forensic odontology: Principles and practice. First edition. London: Academic Press. p 145–171.

Li G, Shi J. 2010. On comparing three artificial neural networks for wind speed forecasting. Appl

Energy 87:2313–2320.

Liversidge HM, Buckberry J, Marquez-Grant N. 2015. Age estimation. Ann Hum Biol 42:297–299.

Liversidge HM, Molleson T. 2004. Variation in crown and root formation and eruption of human deciduous teeth. Am J Phys Anthropol 123:172–180.

Loeser R. 2010. Age-Related Changes in the Mus. Clin Geriatr Med 26:371–386.

Loevy HT, Aduss H, Rosenthal IM. 1987. Tooth eruption and craniofacial development in congenital hypothyroidism: report of case. J Am Dent Assoc 115:429–431.

López-Alcaraz M, González PMG, Aguilera IA, López MB. 2015. Image analysis of pubic bone for age estimation in a computed tomography sample. Int J Legal Med 129:335–346.

Lottering N, Alston-Knox C, MacGregor DM, Izatt M, Grant C, Adam C, Gregory L. 2017. Apophyseal ossification of the iliac crest in forensic age estimation : Computed tomography standards for modern Australian Subadults. J Forensic Sci 62:292–307.

Lottering N, MacGregor DM, Meredith M, Alston CL, Gregory LS. 2013. Evaluation of the Suchey-Brooks method of age estimation in an Australian subpopulation using computed tomography of the pubic symphyseal surface. Am J Phys Anthropol 150:386–399.

Lovejoy CO, Meindl RS, Mensforth RP, Barton TJ. 1985a. Multifactorial determination of skeletal age at death: A method and blind tests of its accuracy. Am J Phys Anthropol 68:1–14.

Lovejoy CO, Meindl RS, Pryzbeck TR, Mensforth RP. 1985b. Chronological metamorphosis of the auricular surface of the ilium: a new method for the determination of adult skeletal age at death. Am J Phys Anthropol 68:15–28.

Lovejoy CO, Meindl RS, Tague R, Latimer B. 1995. The senescent biology of the hominoid pelvis: Its bearing on the pubic symphysis and auricular surface as age-at-death indicators in the human skeleton. Riv di Antropol 73:31–49.

Lucy D, Aykroyd RG, Pollard AM, Solheim T. 1996. A Bayesian approach to adult human age estimation from dental observations by Johanson's age changes. J Forensic Sci 41:189–194.

Luna LH, Aranda CM, Santos AL. 2017. New method for sex prediction using the human non-adult auricular surface of the ilium in the Collection of Identified Skeletons of the University of Coimbra. Int J Osteoarchaeol 27:898–911.

Maggiano IS, Maggiano CM, Clement JG, Thomas CDL, Carter Y, Cooper DML. 2016. Three-dimensional reconstruction of Haversian systems in human cortical bone using synchrotron radiation-based micro-CT: Morphology and quantification of branching and transverse connections across age. J Anat 228:719–732.

Mahato N. 2016. Sacroiliac joints. In: Tubbs R, Shoja M, Loukas M, editors. Bergman's Comprehensive Encyclopedia of Human Anatomic Variation. First edition. Hoboken: John Wiley & Sons. p 165–175.

Mann R, Jantz R, Bass W, Willey P. 1991. Maxillary suture obliteration: a visual method for

estimating skeletal age. J Forensic Sci 36:781–791.

Marciani RD. 2007. Third Molar Removal: An overview of indications, imaging, evaluation, and assessment of risk. Oral Maxillofac Surg Clin North Am 19:1–13.

Márquez-Grant N. 2015. An overview of age estimation in forensic anthropology: perspectives and practical considerations. Ann Hum Biol 42:308–322.

Martins R, Oliveira PE, Schmitt A. 2012. Estimation of age at death from the pubic symphysis and the auricular surface of the ilium using a smoothing procedure. Forensic Sci Int 219:287.e1-e7.

Martrille L, Ubelaker DH, Cattaneo C, Seguret F, Tremblay M, Baccino E. 2007. Comparison of four skeletal methods for the estimation of age at death on white and black adults. J Forensic Sci 52:302–307.

Masset C. 1990. Où en est la paléodémographie ? Bull Mem Soc Anthropol Paris 2:109–121.

Masters P, Bada J, Zigler S. 1977. Aspartic acid racemisation in the human lens during ageing and in cataract formation. Nature 268:71–73.

Mays S. 2012. An investigation of age-related changes at the acetabulum in 18th-19th century ad adult skeletons from Christ Church Spitalfields, London. Am J Phys Anthropol 149:485–492.

Mays S. 2014. A test of a recently devised method of estimating skeletal age at death using features of the adult acetabulum. J Forensic Sci 59:184–187.

Mays S. 2015. The effect of factors other than age upon skeletal age indicators in the adult. Ann Hum Biol 42:332–341.

McCormick W. 1980. Mineralization of the costal cartilages as an indicator of age: Preliminary observations. J Forensic Sci 25:736–741.

McKern TW, Stewart TD. 1957. Skeletal age changes in young American males, Technical report EP 45, Quartermaster Research and Development Command, Natick.

Meindl RS, Lovejoy CO. 1985. ectocranial suture closure: a revised method for the determination of skeletal age at death based on the lateral-anterior sutures. Am J Phys Anthropol 68:57–66.

Meindl RS, Lovejoy CO, Mensforth RP, Walker RA. 1985. A revised method of age determination using the os pubis, with a review and tests of accuracy of other current methods of pubic symphyseal aging. Am J Phys Anthropol 68:29–45.

Mensforth RP, Latimer BM. 1989. Hamann-Todd collection aging studies: Osteoporosis fracture syndrome. Am J Phys Anthropol 80:461–479.

Merritt CE. 2015. The influence of body size on adult skeletal age estimation methods. Am J Phys Anthropol 156:35–57.

Merritt CE. 2017a. Body size as a factor in skeletal age estimation: When size matters and how to deal with it. In: 86th Annual Meeting of the American association of physical anthropologists. New Orleans, p 286.

Merritt CE. 2017b. Inaccuracy and bias in adult skeletal age estimation: Assessing the reliability of eight methods on individuals of varying body sizes. Forensic Sci Int 275:315.e1-315.e11.

Merritt CE. 2018a. Part II – Adult skeletal age estimation using CT scans of cadavers : Revision of the pubic symphysis methods. J Forensic Radiol Imaging 14:50–57.

Merritt CE. 2018b. Part III – Adult skeletal age estimation using CT scans of cadavers : Revision of the auricular surface methods. J Forensic Radiol Imaging 14:58–64.

Merwin DR, Harris EF. 1998. Sibling similarities in the tempo of human tooth mineralization. Arch Oral Biol 43:205–210.

Mesteková S, Průžek J, Velemínska J, Chaumoître K. 2015. A Test of the DSP sexing method on CT images from a modern French sample. J Forensic Sci 60:1295–1299.

Milner GR, Boldsen JL. 2012a. Skeletal age estimation: Where we are and where we should go. In: Dirkmaat D, editor. A Companion to Forensic Anthropology. First edition. London: Wiley-Blackwell. p 224–238.

Milner GR, Boldsen JL. 2012b. Transition analysis: A validation study with known-age modern American skeletons. Am J Phys Anthropol 148:98–110.

Milner GR, Wood JW, Boldsen JL. 2018. Paleodemography: problems, progress, and potential. In: Katzenberg M, Grauer A, editors. Biological Anthropology of the Human Skeleton. Third edition. York: John Wiley & Sons. p 593–633.

Mincer HH, Harris EF, Berryman HE. 1993. The A.B.F.O. Study of third molar development and its use as an estimator of chronological age. J Forensic Sci 38:379–390.

Minegishi S, Ohtani S, Noritake K, Funakoshi T, Ishii N, Utsuno H, Sakuma A, Saitoh H, Yamaguchi S, Marukawa E, Harada H, Uemura K, Sakurada K. 2019. Preparation of dentin standard samples for age estimation based on increased aspartic acid racemization rate by heating. Leg Med 38:25–31.

Miranker M. 2016. A comparison of different age estimation methods of the adult pelvis. J Forensic Leg Med 61:1173–1179.

Miranker M. 2017. Aging using adult human pelvis morphology: effect of occupation or statistical method. In: 86th Annual Meeting of the American association of physical anthropologists. New Orleans, p 288.

Mobasheri A, Batt M. 2016. An update on the pathophysiology of osteoarthritis. Ann Phys Rehabil Med 59:333–339.

Molinari L, Gasser T, Largo RH. 2004. TW3 bone age: RUS/CB and gender differences of percentiles for score and score increments. Ann Hum Biol 31:421–435.

Moorrees CFA, Fanning EA, Hunt EE. 1963a. Age variation of formation stages for ten permanent teeth. J Dent Res 42:1490–1502.

Moorrees CFA, Fanning EA, Hunt EE. 1963b. Formation and resorption of three deciduous teeth in

children. Am J Phys Anthropol 21:205–213.

Mulhern DM, Jones EB. 2005. Test of revised method of age estimation from the auricular surface of the ilium. Am J Phys Anthropol 126:61–65.

Mullins RA, Albanese J. 2018. Estimating biological characteristics with virtual laser data. J Forensic Sci 63:815–823.

Murail P, Brůžek J, Houët F, Cunha E. 2005. DSP: A tool for probabilistic sex diagnosis using worldwide variability in hip-bone measurements. Bull Mem Soc Anthropol Paris 17:167–176.

Murphy RE, Garvin HM. 2018. A morphometric outline analysis of ancestry and sex differences in cranial shape. J Forensic Sci 63:1001–1009.

Musilová B, Dupej J, Brůžek J, Bejdová Š, Velemínská J. 2019. Sex and ancestry related differences between two Central European populations determined using exocranial meshes. Forensic Sci Int 297:364–369.

Navega D, Coelho C, Vicente R, Ferreira MT, Wasterlain S, Cunha E. 2015a. AncesTrees: Ancestry estimation with randomized decision trees. Int J Legal Med 129:1145–1153.

Navega D, Coelho J d. O, Cunha E, Curate F. 2018a. DXAGE: A new method for age at death estimation based on femoral bone mineral density and artificial neural networks. J Forensic Sci 63:497–503.

Navega D, Godinho M, Cunha E, Ferreira MT. 2018b. A test and analysis of Calce (2012) method for skeletal age-at-death estimation using the acetabulum in a modern skeletal sample. Int J Legal Med 132:1447–1455.

Navega D, Vicente R, Vieira DN, Ross AH, Cunha E. 2015b. Sex estimation from the tarsal bones in a Portuguese sample: a machine learning approach. Int J Legal Med 129:651–659.

Nawrocki SP. 2010. The nature and sources of error in the estimation of age at death from the skeleton. In: Age Estimation of the Human Skeleton. First edition. Springfield: Charles C. Thomas. p 79–101.

Nikita E. 2017a. Sex and Ancestry Assessment. In: Osteoarchaeology: A Guide to the macroscopic study of human skeletal remains. First edition. London: Academic Press. p 105–134.

Nikita E. 2017b. Growth Patterns. In: Osteoarchaeology: A Guide to the Macroscopic Study of Human Skeletal Remains. First edition. London: Academic Press. p 243–267.

Nikita E. 2017c. Age estimation. In: Osteoarchaeology A Guide to the Macroscopic Study of Human Skeletal Remains. First edition. London: Academic Press. p 135–173.

Nikita E. 2017d. Osteoarchaeology: A guide to the macroscopic study of human skeletal remains. First edition. London: Academic Press.

Nikita E, Nikitas P. 2019. Skeletal age-at-death estimation: Bayesian versus regression methods. Forensic Sci Int 297:56–64.

Niskanen M, Maijanen H, McCarthy D, Junno JA. 2013. Application of the anatomical method to

estimate the maximum adult stature and the age-at-death stature. Am J Phys Anthropol 152:96–106.

Obertová Z, Cattaneo C. 2018. Child trafficking and the European migration crisis: The role of forensic practitioners. Forensic Sci Int 282:46–59.

Oettle A, Steyn M. 2000. Age estimation from sternal ends of ribs by phase analysis in South African Blacks. J Forensic Sci 45:1071–1079.

Ohtake PJ. 2008. The impact of obesity on walking: Implications for fitness assessment and exercise prescription. Cardiopulm Phys Ther J 19:52–53.

Ohtani S. 1998. Rate of aspartic acid racemization in bone. Am J Forensic Med Pathol 19:284–287.

Ohtani S, Yamamoto T. 2010. Age estimation by amino acid racemization in human teeth. J Forensic Sci 55:1630–1633.

Olze A, Schmeling A, Taniguchi M, Maeda H, Van Niekerk P, Wernecke KD, Geserick G. 2004. Forensic age estimation in living subjects: The ethnic factor in wisdom tooth mineralization. Int J Legal Med 118:170–173.

Osborne DL, Simmons TL, Nawrocki SP. 2004. Reconsidering the auricular surface as an indicator of age at death. J Forensic Sci 49:1–7.

Ousley SD, Hollinger R. 2012. The Pervasiveness of Daubert. In: Dirkmaat DC, editor. A Companion to Forensic Anthropology. First edition. London: Wiley-Blackwell. p 654–665.

Pattamapaspong N, Kanthawang T, Singsuwan P, Sansiri W, Prasitwattanaseree S, Mahakkanukrauh P. 2019. Efficacy of three-dimensional cinematic rendering computed tomography images in visualizing features related to age estimation in pelvic bones. Forensic Sci Int 294:48–56.

Perreard Lopreno G. 2007. Adaptation structurelle des os du membre supérieur et de la clavicule à l'activité. Dissertation dissertation, University of Geneva.

Phenice T. 1969. A newly developed visual method of sexing the os pubis. Am J Phys Anthropol 30:297–301.

Poláček L. 2008. Das Hinterland des frühmitelalterlichen Zentrums in Mikulčice. Stand und Perspektiven der Forschung. In: Poláček L, editor. Das wirtschäftliche Hinterland der frühmittelalterlichen Zentren. Internationale Tagungen in Mikulčice VI. Brno: Archeologický ústav Akademie věd ČR. p 257–298.

Powers R. 1962. The disparity between known age and age as estimated by cranial suture closure. Man 62:52–54.

Prentice A, Schoenmakers I, Laskey AM, de Bono S, Ginty F, Goldberg GR. 2006. Nutrition and bone growth and development. Proc Nutr Soc 65:348–360.

Puhakka KB, Melsen F, Jurik AG, Boel LW, Vesterby A, Egund N. 2004. MR imaging of the normal sacroiliac joint with correlation to histology. Skeletal Radiol 33:15–28.

Purves S, Woodley L, Hackman L. 2011. Age determination in the adult. In: Black S, Ferguson E, editors. Forensic anthropology: 2000 to 2010. First edition. London: CRC Press. p 29–59.

Ramsthaler F, Kettner M, Gehl A, Verhoff M a. 2010. Digital forensic osteology: Morphological sexing of skeletal remains using volume-rendered cranial CT scans. Forensic Sci Int 195:148–152.

Raxter MH, Auerbach BM, Ruff CB. 2006. Revision of the fully technique for estimating statures. Am J Phys Anthropol 130:374–384.

Raxter MH, Ruff CB. 2018. Full Skeleton Stature Estimation. In: Latham K, Bartelink E, Finnegan M, editors. New Perspectives in Forensic Human Skeletal Identification. First edition. London: Academic Press. p 105–113.

Renz H, Radlanski RJ. 2006. Incremental lines in root cementum of human teeth - A reliable age marker? HOMO 57:29–50.

Reppien K, Sejrsen B, Lynnerup N. 2006. Evaluation of post-mortem estimated dental age versus real age: A retrospective 21-year survey. Forensic Sci Int 159:84–88.

Rissech C, Appleby J, Cosso A, Reina F, Carrera A, Thomas R. 2018. The influence of bone loss on the three adult age markers of the innominate. Int J Legal Med 132:289–300.

Rissech C, Estabrook GF, Cunha E, Malgosa A. 2006. Using the acetabulum to estimate age at death of adult males. J Forensic Sci 51:213–229.

Rissech C, Estabrook GF, Cunha E, Malgosa A. 2007. Estimation of age-at-death for adult males using the acetabulum, applied to four Western European populations. J Forensic Sci 52:774–778.

Rissech C, Schmitt A, Malgosa A, Cunha E. 2004. Influencia de las patologías en los indicadores de edad adulta del coxal: estudio preliminar. Antropol Port 20:267–279.

Rissech C, Winburn AP, San-Millán M, Sastre J, Rocha J. 2019. The acetabulum as an adult age marker and the new IDADE2 (the IDADE2 web page). Am J Phys Anthropol 169:757–764.

Risser JC. 1958. The Iliac apophysis; an invaluable sign in the management of scoliosis. Clin Orthop Relat Res 11:111–119.

Ritz-Timme S, Cattaneo C, Collins MJ, Waite ER, Schütz HW, Kaatsch HJ, Borrman HI. 2000. Age estimation: The state of the art in relation to the specific demands of forensic practise. Int J Legal Med 113:129–136.

Rivera-Sandoval J, Monsalve T, Cattaneo C. 2018. A test of four innominate bone age assessment methods in a modern skeletal collection from Medellin, Colombia. Forensic Sci Int 282:232.e1-232.e8.

Robling A, Stout S. 2007. Histomorphometry of Human Cortical Bone: Applications to Age Estimation. In: Katzenberg M, Saunders S, editors. Biological Anthropology of the Human Skeleton. Second Edition. New York: Wiley-Liss. p 149–182.

Rogers J, Watt I, Dieppe P. 1990. Comparison of visual and radiographic detection of bony changes at the knee joint. Br Med J 300:367–368.

Roksandic M, Vlak D, Schillaci MA, Voicu D. 2009. Technical note: Applicability of tooth cementum annulation to an archaeological population. Am J Phys Anthropol 140:583–588.

Rösing FW, Graw M, Marré B, Ritz-Timme S, Rothschild M a, Rötzscher K, Schmeling A, Schröder I, Geserick G. 2007. Recommendations for the forensic diagnosis of sex and age from skeletons. Homo 58:75–89.

Rougé-Maillart C, Jousset N, Vielle B, Gaudin A, Telmon N. 2007. Contribution of the study of acetabulum for the estimation of adult subjects. Forensic Sci Int:103–110.

Rougé-Maillart C, Telmon N, Rissech C, Malgosa A, Rougé D. 2004. The determination of male adult age at death by central and posterior coxal analysis—A preliminary study. J Forensic Sci 49:1–7.

Rougé-Maillart C, Vielle B, Jousset N, Chappard D, Telmon N, Cunha E. 2009. Development of a method to estimate skeletal age at death in adults using the acetabulum and the auricular surface on a Portuguese population. Forensic Sci Int 188:91–95.

Rozkovcová E, Marková M, Mrklas L. 2005. Nové přístupy k problematice třetího moláru. Česká Stomatol 105:119–128.

Ruengdit S, Prasitwattanaseree S, Mekjaidee K, Sinthubua A, Mahakkanukrauh P. 2018. Age estimation approaches using cranial suture closure: A validation study on a Thai population. J Forensic Leg Med 53:79–86.

Sakaue K. 2006. Application of the Suchey–Brooks system of pubic age estimation to recent Japanese skeletal material. Anthropol Sci 114:59–64.

San-Millán M, Rissech C, Turbón D. 2017. New approach to age estimation of male and female adult skeletons based on the morphological characteristics of the acetabulum. Int J Legal Med 131:501–525.

San-Millán M, Rissech C, Turbón D. 2019. Application of the recent SanMillán–Rissech acetabular adult aging method in a North American sample. Int J Legal Med 133:909–920.

Santos F, Guyomarc'h P, Bruzek J. 2014. Statistical sex determination from craniometrics: Comparison of linear discriminant analysis, logistic regression, and support vector machines. Forensic Sci Int 245:204.e1-e8.

Sarkar A, Singh M, Bansal N, Kapoor S. 2011. Effects of obesity on balance and gait alterations in young adults. Indian J Physiol Pharmacol 55:227–233.

Saunders SR, Fitzgerald C, Rogers T, Dudar C, McKillop H. 1992. A test of several methods of skeletal age estimation using a documented archaeological sample. Can Soc Forensic Sci J 25:97–118.

Savall F, Rérolle C, Hérin F, Dédouit F, Rougé D, Telmon N, Saint-Martin P. 2016. Reliability of

the Suchey-Brooks method for a French contemporary population. Forensic Sci Int 266:586-e1.

Schaefer M, Black S, Scheuer L. 2009. Juvenile Osteology: A Laboratory and Field Manual. First edition. London: Academic Press.

Schaefer M, Geske N, Cunningham CA. 2018. A decade of development in juvenile aging. In: Latham K, Bartelink E, Finnegan M, editors. New perspectives in forensic human skeletal identification. First edition. Academic Press. p 45–60.

Schmeling A, Schulz R, Danner B, Rösing FW. 2006. The impact of economic progress and modernization in medicine on the ossification of hand and wrist. Int J Legal Med 120:121–126.

Schmeling A, Schulz R, Reisinger W, Mühler M, Wernecke KD, Geserick G. 2004. Studies on the time frame for ossification of the medial clavicular epiphyseal cartilage in conventional radiography. Int J Legal Med 118:5–8.

Schmidt S, Ottow C, Pfeiffer H, Heindel W, Vieth V, Schmeling A, Schulz R. 2017. Magnetic resonance imaging-based evaluation of ossification of the medial clavicular epiphysis in forensic age assessment. Int J Legal Med 131:1665–1673.

Schmitt A. 2001. Variabilité de la sénescence du squelette humain. Réflexions sur les indicateurs de l'âge au décès: à la recherche d'un outil performant. Dissertation thesis. University of Bordeaux.

Schmitt A. 2004. Age-at-death assessment using the os pubis and the auricular surface of the ilium: A test on an identified Asian sample. Int J Osteoarchaeol 14:1–6.

Schmitt A. 2005. Une nouvelle méthode pour estimer l'âge au décès des adultes à partir de la surface sacro-pelvienne iliaque. Bull Mem Soc Anthropol Paris 17:89–101.

Schmitt A, Murail P, Cunha E, Rougé D. 2002. Variability of the pattern of aging on the human skeleton: Evidence from bone indicators and implication on age at death estimation. J Forensic Sci 47:1203–1209.

Schour I, Massler M. 1941. The development of the human dentition. J Am Dent Assoc 28:1153–1160.

Scott G, Navega D, Coelho J, Cunha E, Irish J. 2016. rASUDAS: A new method for estimating ancestry from tooth crown and root morphology. In: 85th Annual Meeting of the American association of physical anthropologists. Atlanta. p 285.

Scott G, Pilloud M, Navega D, D'Oliveira J, Cunha E, Irish J. 2018. rASUDAS: A new web-based application for estimating ancestry from tooth morphology. Simulation 1:18–31.

Seguchi N, Dudzik B. 2019. 3D Data Acquisition for Bioarchaeology, Forensic Anthropology, and Archaeology. First edition. London: Academic Press.

Seguchi N, Dudzik B, Murphy M, Prentiss A. 2019. Introduction. In: Seguchi N, Dudzik B, editors.

3D Data Acquisition for Bioarchaeology, Forensic Anthropology, and Archaeology. First edition. London: Academic Press. p 1–16.

Shirley NR, Jantz RL. 2011. Spheno-occipital synchondrosis fusion in modern Americans. J Forensic Sci 56:580–585.

Shirley NR, Ramirez Montes PA. 2015. Age estimation in forensic anthropology: Quantification of observer error in phase versus component-based methods. J Forensic Sci 60:107–111.

Slice DE, Algee-Hewitt BFB. 2015. Modeling Bone Surface Morphology: A fully quantitative method for age-at-death estimation using the pubic symphysis. J Forensic Sci 60:835–843.

Smith HB. 1991. Standards of human tooth formation and dental age assessment. In: Kelly M, Larsen C, editors. Advances in dental anthropology. First edition. New York: Wiley-Liss. p 143–168.

Smith SL. 2007. Stature estimation of 3-10-year-old children from long bone lengths. J Forensic Sci 52:538–546.

Solheim T, Vonen A. 2006. Dental age estimation, quality assurance and age estimation of asylum seekers in Norway. Forensic Sci Int 159:56–60.

Spradley MK, Jantz RL. 2011. Sex Estimation in Forensic Anthropology : Skull Versus Postcranial Elements. J Forensic Sci 56:289–296.

Spradley MK, Jantz RL, Robinson A, Peccerelli F. 2008. Demographic change and forensic identification: problems in metric identification of Hispanic skeletons. 53:21–28.

Spradley MK, Weisensee K. 2017. Ancestry Estimation. The Importance, The History, and The Practice. In: Langley NR, Tersigni-Tarrant MA, editors. Forensic Anthropology: A Comprehensive Introduction. Second edition. Boca Raton: CRC Press. p 163–174.

Standring S. 2016. Gray's Anatomy: the Anatomical Basis of Clinical Practice. Forty first edition. London: Elsevier.

Steadman DW, Adams BJ, Konigsberg LW. 2006. Statistical basis for positive identification in forensic anthropology. Am J Phys Anthropol 131:15–26.

Stoyanova D, Algee-Hewitt BFB, Slice DE. 2015. An enhanced computational method for age-at-death estimation based on the pubic symphysis using 3D laser scans and thin plate splines. Am J Phys Anthropol 158:431–440.

Stoyanova DK, Algee-Hewitt BFB, Kim J, Slice DE. 2017. A computational framework for age-at-death estimation from the skeleton: surface and outline analysis of 3D laser scans of the adult pubic symphysis. J Forensic Sci 62:1434–1444.

Stull KE, L'Abbé EN, Ousley SD. 2017. Subadult sex estimation from diaphyseal dimensions. Am J Phys Anthropol 163:64–74.

Techataweewan N, Panthongviriyakul C, Toomsan Y, Mothong W, Kanla P, Amarttayakong P, Tayles N. 2017a. The body donation in Thailand: donors at Khon Kaen University. Ann

Anatomy-Anatomischer Anzeiger 216:142–151.

Techataweewan N, Tuamsuk P, Toomsan Y, Woraputtaporn W, Prachaney P, Tayles N. 2017b. A large modern Southeast Asian human skeletal collection from Thailand. Forensic Sci Int 278:406.e1-406.e6.

Telmon N, Gaston A, Chemla P, Blanc A, Joffre F, Rougé D. 2005. Application of the Suchey - Brooks method to three - dimensional imaging of the pubic symphysis. J Forensic Sci 50:1–6.

Thesleff I, Sharpe P. 1997. Signalling networks regulating dental development. Mech Dev 67:111–123.

Todd TW, Lyon DW. 1924. Endocranial suture closure: its progress and age relationship. Part I. Adult males of white stock. Am J Phys Anthropol 7:325–384.

Todd TW, Lyon DW. 1925. Suture closure: Its progress and age relationship. Part IV. Ectocranial closure in adult males of Negro stock. Am J Phys Anthropol 8:149–168.

Todd WT. 1920. Age changes in the pubic bone. I. The male white pubis. Am J Phys Anthropol 3:285–334.

Traithepchanapai P, Mahakkanukrauh P, Kranioti EF. 2016. History, research and practice of forensic anthropology in Thailand. Forensic Sci Int 261:167.e1-167.e6.

Tsuji A, Ishiko A, Takasaki T, Ikeda N. 2002. Estimating age of humans based on telomere shortening. Forensic Sci Int 126:197–199.

Ubelaker DH. 1978. Estimating age at death from human skeletons: An overview. J Forensic Sci 32:1254–63.

Ubelaker DH. 1989. Human Skeletal Remains: Excavation, Analysis, Interpretation. Washington DC: Taraxacum.

Ubelaker DH. 2018a. Estimation of immature age from the dentition. In: Latham K, Bartelink E, Finnegan M, editors. New perspectives in forensic human skeletal identification. First edition. London: Academic Press. p 61–64.

Ubelaker DH. 2018b. Forensic Anthropology: Methodology and Applications. In: Katzenberg N, Grauer A, editors. Biological Anthropology of the Human Skeleton. Third Edition New York: John Wiley & Sons. p 43–71.

Ubelaker DH, DeGaglia CM. 2017. Population variation in skeletal sexual dimorphism. Forensic Sci Int 278:407.e1-407.e7.

Ubelaker DH, Khosrowshahi H. 2019. Estimation of age in forensic anthropology : historical perspective and recent methodological advances recent methodological advances. Forensic Sci Res 4:1–9.

Verhoff MA, Ramsthaler F, Krähahn J, Deml U, Gille RJ, Grabherr S, Thali MJ, Kreutz K. 2008. Digital forensic osteology-Possibilities in cooperation with the Virtopsy®project. Forensic Sci Int 174:152–156.

Viggiano D, Thanassoulas T, Di-Cesare C, Cacciola G, Giorgio NM, Pitsios T, Passiatore C. 2015. A low-cost system to acquire 3D surface data from anatomical samples. Eur J Anat 19:343–349.

Villa C, Buckberry J, Cattaneo C, Frohlich B, Lynnerup N. 2015a. Quantitative analysis of the morphological changes of the pubic symphyseal face and the auricular surface and implications for age at death estimation. J Forensic Sci 60:556–565.

Villa C, Buckberry J, Cattaneo C, Lynnerup N. 2013a. Technical note: Reliability of suchey-brooks and buckberry-chamberlain methods on 3D visualizations from CT and laser scans. Am J Phys Anthropol 151:158–163.

Villa C, Gaudio D, Cattaneo C, Buckberry J, Wilson AS, Lynnerup N. 2015b. Surface Curvature of Pelvic Joints from Three Laser Scanners: Separating Anatomy from Measurement Error. J Forensic Sci 60:374–381.

Villa C, Hansen MN, Buckberry J, Cattaneo C, Lynnerup N. 2013b. Forensic age estimation based on the trabecular bone changes of the pelvic bone using post-mortem CT. Forensic Sci Int 233:393–402.

Walker PL. 2008. Sexing skulls using discriminant function analysis of visually assessed traits. Am J Phys Anthropol 136:39–50.

Ward R. 2016. Hip joint. In: Tubbs R, Shoja M, Loukas M, editors. Bergman's Comprehensive Encyclopedia of Human Anatomic Variation. First edition. Hoboken: John Wiley & Sons. p 176–180.

Wescott DJ, Drew JL. 2015. Effect of obesity on the reliability of age-at-death indicators of the pelvis. Am J Phys Anthropol 156:595–605.

Willems G, Olmen A Van, Spiessens B, Carels C. 2001. Dental Age Estimation in Belgian Children: Demirjian's Technique Revisited. J Forensic Sci 46:893–895.

Willershausen B, Löffler N, Schulze R. 2001. Analysis of 1202 orthopantograms to evaluate the potential of forensic age determination based on third molar developmental stages. Eur J Med Res 6:377–84.

Willmott CJ, Matsuura K. 2005. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. Clim Res 30:79–82.

Winburn AP. 2018. Validation of the acetabulum as a skeletal indicator of age at death in modern European-Americans. J Forensic Sci 64:989–1003.

Winburn AP, Stock MK. 2019. Reconsidering osteoarthritis as a skeletal age indicator. In: 88th Annual Meeting of American association of physical anthropologists. Cleveland.

Wink AE. 2014. Pubic symphyseal age estimation from three-dimensional reconstructions of pelvic ct scans of live individuals. J Forensic Sci 59:696–702.

Wittschieber D, Schmeling A, Schmidt S, Heindel W, Pfeiffer H, Vieth V. 2013a. The Risser sign

for forensic age estimation in living individuals: A study of 643 pelvic radiographs. Forensic Sci Med Pathol 9:36–43.

Wittschieber D, Vieth V, Domnick C, Pfeiffer H, Schmeling A. 2013b. The iliac crest in forensic age diagnostics: Evaluation of the apophyseal ossification in conventional radiography. Int J Legal Med 127:473–479.

Wittwer-Backofen U, Gampe J, Vaupel JW. 2004. Tooth Cementum Annulation for Age Estimation: Results from a Large Known-Age Validation Study. Am J Phys Anthropol 123:119–129.

Wood C. 2015. The age-related emergence of cranial morphological variation. Forensic Sci Int 251:220.e1-e20.

Xanthopoulou P, Valakos E, Youlatos D, Nikita E. 2018. Assessing the accuracy of cranial and pelvic ageing methods on human skeletal remains from a modern Greek assemblage. Forensic Sci Int 286:266.e1-266.e8.

Zapico S, DeGaglia CM, Adserias-Garriga J. 2019a. Age estimation based on chemical approaches. In: Adserias-Garriga J, editor. Age Estimation: A Multidisciplinary Approach. First edition. London: Academic Press. p 199–211.

Zapico S, Thomas C, Zoppis S. 2019b. Age estimation based on molecular biology approaches. In: Age Estimation: A Multidisciplinary Approach. First edition. London: Academic Press. p 213–223.

Zapico S, Ubelaker DH. 2013. Applications of physiological bases of ageing to forensic sciences. Estimation of age-at-death. Ageing Res Rev 12:605–617.

Zbieć-Piekarska R, Spólnicka M, Kupiec T, Makowska Z, Spas A, Parys-Proszek A, Kucharczyk K, Płoski R, Branicki W. 2015. Examination of DNA methylation status of the ELOVL2 marker may be useful for human age prediction in forensic science. Forensic Sci Int Genet 14:161–167.

Zinni D, Crowley K. 2017. Appendix A: Application of Dentition in Forensic Anthropology. In: Langley NR, Tersigni-Tarrant MA, editors. Forensic Anthropology: A Comprehensive Introduction. Second edition. Boca Raton: CRC Press. p 365–380.

Web sources

www.atlas.dentistry.qmul.ac.uk

http://www.francecasts.com/

# APPENDICES

## A. SUBMITTED PUBLICATIONS

**Kotěrová, A.**, Velemínská, J., Dupej, J., Brzobohatá, H., Pilný, A., Brůžek, J. 2017. Disregarding population specificity: its influence on the sex assessment methods from the tibia. *International Journal of Legal Medicine*. 131.1: 251–261.

**Kotěrová, A.**, Navega, D., Štěpanovský, M., Buk, Z., Brůžek, J., Cunha, E. 2018. Age estimation of adult human remains from hip bones using advanced methods. *Forensic Science International*. 287:163–175.

**Kotěrová, A.**, Králík, V., Rmoutilová, R., Friedl, L., Růžička, P., Velemínská, J., Marchal F., Brůžek, J. 2019. Impact of 3D Surface Scanning Protocols on the Os Coxae Digital Data: Implications for Sex and Age-at-death Assessment. *Journal of forensic and legal medicine*. 68: 101866.

**Kotěrová, A.**, Velemínská, J., Cunha, E., Brůžek, J. 2018. A validation study of the Stoyanova et al. method (2017) for age-at-death estimation quantifying the 3D pubic symphyseal surface of adult males of European populations. *International Journal of Legal Medicine.* 133:603–612.

## Appendix A.1

Int J Legal Med (2017) 131:251–261
DOI 10.1007/s00414-016-1413-5

ORIGINAL ARTICLE

# Disregarding population specificity: its influence on the sex assessment methods from the tibia

**Anežka Kotěrová[1] · Jana Velemínská[1] · Ján Dupej[2] · Hana Brzobohatá[3] · Aleš Pilný[4] · Jaroslav Brůžek[1]**

Received: 27 January 2016 / Accepted: 30 June 2016 / Published online: 20 July 2016
© Springer-Verlag Berlin Heidelberg 2016

**Abstract** Forensic anthropology has developed classification techniques for sex estimation of unknown skeletal remains, for example population-specific discriminant function analyses. These methods were designed for populations that lived mostly in the late nineteenth and twentieth centuries. Their level of reliability or misclassification is important for practical use in today's forensic practice; it is, however, unknown. We addressed the question of what the likelihood of errors would be if population specificity of discriminant functions of the tibia were disregarded. Moreover, five classification functions in a Czech sample were proposed (accuracies 82.1–87.5 %, sex bias ranged from −1.3 to −5.4 %). We measured ten variables traditionally used for sex assessment of the tibia on a sample of 30 male and 26 female models from recent Czech population. To estimate the classification accuracy and error (misclassification) rates ignoring population specificity, we selected published classification functions of tibia for the Portuguese, south European, and the North American populations. These functions were applied on the dimensions of the Czech population. Comparing the classification success of the reference and the tested Czech sample showed that females from Czech population were significantly overestimated and mostly misclassified as males. Overall accuracy of sex assessment significantly decreased (53.6–69.7 %), sex bias −29.4–100 %, which is most probably caused by secular trend and the generally high variability of body size. Results indicate that the discriminant functions, developed for skeletal series representing geographically and chronologically diverse populations, are not applicable in current forensic investigations. Finally, implications and recommendations for future research are discussed.

**Keywords** Forensic anthropology population data · Sex determination · Tibia · Population specificity · Discriminant function analysis · GAME method

✉ Anežka Kotěrová
koterova.a@seznam.cz; koterova@natur.cuni.cz

[1] Department of Anthropology and Human Genetics, Faculty of Science, Charles University in Prague, Viničná 7, Praha 2, 128 43 Prague, Czech Republic

[2] Department of Software and Computer Science, Faculty of Mathematics and Physics, Charles University in Prague, Malostranské nám. 25, 118 00 Prague, Czech Republic

[3] Department of Prehistory, Institute of Archaeology of the Academy of Science, Letenská 4, 118 01 Praha, Czech Republic

[4] Department of Computer Science, Faculty of Electrical Engineering, Czech Technical University in Prague, Karlovo náměstí 13, 121 35 Prague 2, Czech Republic

## Introduction

Reliable sex determination is of paramount importance in forensic practice [1–3]. It provides a component of an individual's biological profile complementing age at death estimation, stature, and population affinity assessment [1, 4]. It is widely accepted that the most accurate sex estimation from skeleton can be performed on pelvis and then skull [5, 6]. Unfortunately, they are not always preserved in forensic cases [7] and in such situations, one can use long bones. Especially robust ones like femur or tibia which are often better preserved can be used for sex estimation [8, 9]. Numerous studies were published that prove usefulness of tibia in forensic anthropology [8, 10–21]. Almost all of them agree that its proximal end is the most sexual dimorphic. Generally, the accuracies of sex determination are high (above 85 % when multiple measurements are

used) especially with proximal and distal widths as well as circumference [22].

In sex assessment, discriminant function analysis (DFA) is the most often used classification tool among authors. Their advantages and shortcomings have been discussed recently [23, 24]. DFA tend to be samples-specific and their generalization needs to be assessed in different samples [25].

Any attempts to develop standards for sex estimation of human skeletal remains must take into account that the pattern of sexual dimorphism varies among human populations. Therefore, it is important to avoid the application of metric standards proposed in different populations from different periods of time than the studied sample [26]. Also, as several studies have shown, body size has changed over generations in the population as a consequence of secular trend [27–29]. These changes are especially related to the body height [30]. Secular trend in stature is widely observed in humans since the nineteenth century, and this trend directly impacts the dimensions of long bones [31–34].

Major problems with methods that use size-based variables are that standards can be influenced by secular trend and are usually population-specific. On the other hand, this is to a certain extent also true for shape-based characteristics [22]. Thus, it should be assumed that methods for sex determination, based on skeleton collections of known sex from the first half of the twentieth century, cannot guarantee the same reliability of results when they are used in attempts to identify unknown human remains from recent populations [35].

Application of metric data from one population into the DFA derived from different population groups results in high classification error, and the results are also affected by large sex bias [36, 37]. In a study designed to quantify the effect of applying Euro-American [38] and South African of European ancestry [39] standards to Australian population sample [40], classification accuracy was approximately the same in Australian sample as in target sample (e.g., 80–83 %). On the other hand, correct estimation of sex was distinctly distorted by unacceptable sex bias, i.e., the difference in correct sex determination of males compared to females (which was 31 and 36 %, respectively).

The effect of disregarding population specificity on the classification accuracy of DFA is often mentioned [25] but is rarely described for different parts of the skeleton. In recent publications, we found such studies only for crania from Western Australian and Indian populations [25, 30] and in European populations for the clavicle [41], the calcaneus [37], and for the femur [36]. Recently, many authors have argued for the development and use of population-specific formulae for diverse parts of skeleton when metric data are used [25, 30, 35, 41–45].

The first objective of the present study is to propose classification functions (CFs) for sex estimation in recent Czech population based on CT imaging. The collection of osteometric data from CT images is a reliable and acceptable source and is utilized more and more often for sex assessment in forensic anthropology [46–51]. The other objective is to simulate ignoring population specificity and practically show the range of errors in sex classification.

## Material and methods

### Material

In this study, we used surface models which were constructed from anonymized CT scans from angiography (the slice increment was set at 0.5 mm). Tibial 3D models were reconstructed using specialized software Mimics (Materialise, Leuven, Belgium) and thresholding method. Our virtual material comes from Czech population of the twenty-first century. The age, sex, and race of all specimens are known. Sample consists of nonpathological 56 left human tibiae where 30 belong to male and 26 to female. Male individuals were born between the years 1943 and 1980, and the mean age was 56.1 years (ranging from 31 to 68). Female individuals were born between the years 1920 and 1978 when their mean age was 69.0 years (ranging from 33 to 91). Table 1 shows age distributions, mean ages, and standard deviations. The same material was originally utilized in the study of Brzobohatá et al. [52], where authors performed sex determination from tibiae using geometric morphometrics.

### Selection of measured dimensions

Following relevant published works, we chose ten dimensions commonly measured in anthropology. Their list together with used abbreviations, references where they were defined and landmarks used for computing these dimensions are given in Table 2. Most of the chosen dimensions were defined by Martin and Saller [53] marked by letter M; in other cases, the authors of the definitions is given in the same table.

**Table 1** Characteristics of the Czech sample by age and sex

| Sex | Total | Mean age (year) | SD | 30–40 years | 41–60 years | 61 years and over |
|---|---|---|---|---|---|---|
| Male | 30 | 56.1 | 10.1 | 3 | 13 | 14 |
| Female | 26 | 69.0 | 11.6 | 1 | 3 | 22 |

*SD* standard deviation

**Table 2** Selected and measured dimensions with their name, references, and used landmarks

| Abbreviation | | Main landmarks | Auxiliary landmarks | Name | Reference |
|---|---|---|---|---|---|
| 1 | M3 | 11, 12 | 22, 23, 24[a] | Proximal epiphyseal breadth | Martin and Saller [53] |
| 2 | BB | 9, 10 | | Biarticular breadth | Holland [11] |
| 3 | M3a | 1, 2 | | Medial condyle articular width | Martin and Saller [53] |
| 4 | M3b | 5, 6 | | Medial condyle articular length | Martin and Saller [53] |
| 5 | M4a | 3, 4 | | Lateral condyle articular width | Martin and Saller [53] |
| 6 | M4b | 7, 8 | | Lateral condyle articular length | Martin and Saller [53] |
| 7 | M6 | 17, 21 | 18, 20[b] | Distal epiphysel breadth | Martin and Saller [53] |
| 8 | DB | 17, 19 | | Distal breadth | Işcan and Miller-Shaivitz [10] |
| 9 | M8a | 13, 15 | | Anteroposterior diameter | Martin and Saller [53] |
| 10 | M9a | 14, 16 | | Transverse diameter | Martin and Saller [53] |

[a] Those landmarks were used to construct medial plane and landmarks 11 and 12 lay in the tangent plane which is parallel to defined medial plane

[b] Landmark 21 was computed as the middle point between landmarks 18 and 20

The selection was adjusted to several conditions:

1. Dimensions had to correspond with the selection of discriminant functions.
2. According to the related literature, the most dimorphic dimensions were selected (area of the knee joint mostly).
3. Circumferential dimensions were excluded because of the difficulty of their reproducibility in a virtual environment as opposed to measuring real dry bones.
4. Finally, length was also left out as the length of the bone does not perceptibly contribute to sex classification, compared to joint dimensions.

Measurements were carried out in the Morphome3cs software (www.morphome3cs.com), which was developed in the Department of Software and Computer Science under the Charles University in Prague. In the environment of this program, we located 23 landmarks on surface models of each bone and one landmark (number 21) was computed by the software. For the dimension M3, we first had to construct the medial plane (defined by landmarks 22–24) and then we located landmarks 11 and 12 as the most laterally prominent points of proximal part. These two points lied on the tangent plane which is parallel to defined medial plane. Dimension M6 was measured by locating auxiliary landmarks 18 and 20 and defining landmark 21 as their midpoint. Dimension M6 is the distance between the landmarks 17 and 21. Other dimensions were measured simply as the distances between two relevant landmarks. The locations of all landmarks are documented in Fig. 1. All metric values were collected by one observer.

### Intraobserver error

Intraobserver rate was calculated by relocating all main landmarks of seven randomly selected tibiae. Repeated measurements were performed seven times with at least 1-day interval between them. Average error was calculated in Morphome3cs software and reached 0.4734 mm.

### Selection of published discriminant functions

The process of selection of discriminant functions follows these requirements:

1. Only multidimensional discriminant functions were chosen, because one dimension cannot reliably determine the sex, due to overlapping features of sexual dimorphism between female and male [54].
2. Sex bias could not exceed 7.5 %.
3. The accuracy of sex estimation was greater than 75 %.
4. Selection was made with regard to the criteria and limits of the selected metric dimensions.
5. Selected functions were designed for both European and non-European populations from various periods of time.

Based on these criteria, we chose nine published discriminant functions whose coefficients for each variable, constants, sectioning points, and levels of accuracy are summarized in Table 3. Discriminant functions 1–4 were proposed by Brůžek [14] for a Portuguese population whose birth falls between the early nineteenth and early twentieth centuries. DF 5–8 proposed by Kranioti and Apostol [21] for a sample of south European populations (Spanish, Italian, Greek, and pooled populations) whose death falls into the second half of twentieth century. Last discriminant function DF9 proposed by Işcan and Shaivitz [10] for the North American population whose was born between the early nineteenth century and first half of the twentieth century.

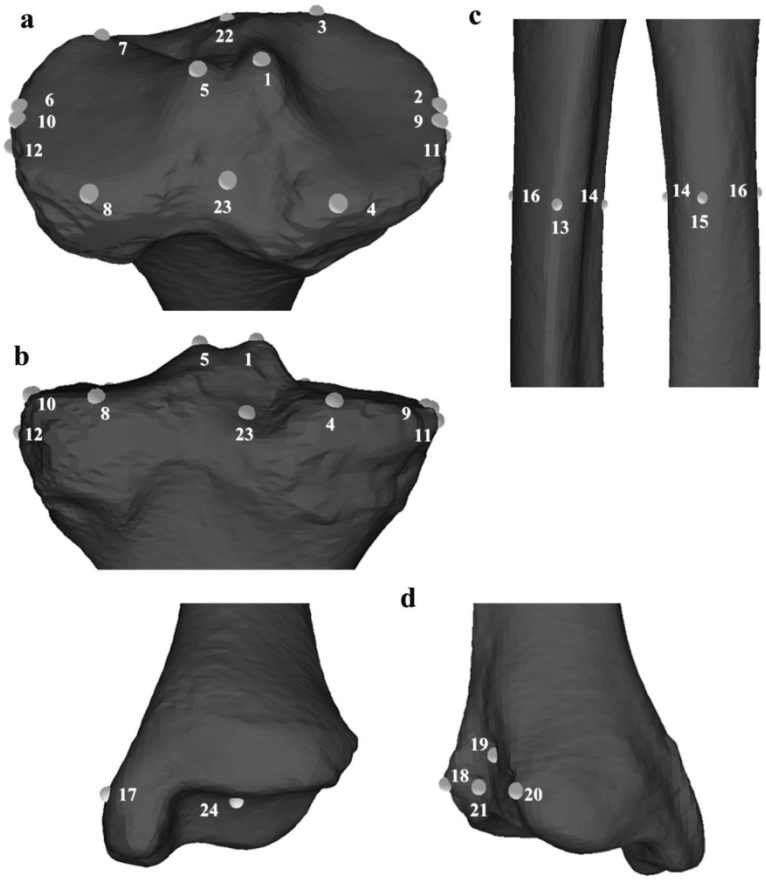


**Fig. 1** Location of landmarks on the **a** proximal part (1–12, 22, 23), **b** proximal part posterior view, **c** body (13–16), and **d** distal part (17–21, 24) of tibia

To evaluate classification performance, accuracy and sex bias were used. According to several authors [25, 55, 56] sex bias is computed as the difference between the classification accuracy of males and that of females, both expressed as percentages. A sex bias lower than 5 % is ideal in forensic anthropology [25]; however, to increase the number of functions available for comparison, a limit of 7.5 % was used.

### Selection of dimensions used for proposing own CF

For classification purposes, the most sexual dimorphic dimensions of tibia were chosen. We thus preferred the dimensions on the proximal part of bone in the area of knee joint (BB, M3, M3a, M3b, M4a, M4b). The last dimension M6 was selected because several authors [9, 12, 14, 20] had used some of the dimensions which measured distal breadth of bone in combination with measurements on the proximal part. Then, using the software Statistica [57], we tested combinations of two or more dimensions and selected the ones with the greatest discriminatory power.

### Statistical analyses

Basic statistical characteristics and the application of Czech dimensions into the selected discriminant functions were performed in MS Excel. Two-sample *t* test was used to quantify the level of dimorphism in the measured lengths. Descriptive statistics for both sexes of the Czech sample, including standard deviation and *t* test for each dimension are shown in Table 4. As expected, males have greater tibial dimensions than females. *t* Test indicates that all of these differences are statistically significant ($\alpha = 0.05$). Only one dimension (DB) was found above the significance level. For all these statistical analyses and for computing our own discriminant functions as well, Statistica software [57] was used. Also tenfold cross-validation was done for linear discriminant functions in program R.

### Classification techniques

For the classification, we used two different approaches. Linear DFA in software Statistica and one advanced data modeling method based on the Group of Adaptive Models

**Table 3** Selection of published discriminant functions with coefficients, sectioning points, and accuracy of sex determination

| | Discriminant functions | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | DF1 Brůžek [14] Portuguese population | DF2 Brůžek [14] | DF3 Brůžek [14] | DF4 Brůžek [14] | DF5 Kranioti and Apostol [21] Spanish, Italian, Greek, and pooled populations | DF6 Kranioti and Apostol [21] | DF7 Kranioti and Apostol [21] | DF8 Kranioti and Apostol [21] | DF9 Işcan and Miller-Shaivitz [10] North American population |
| Variable | Coefficients | | | | | | | | |
| M3 | | | | | 0.2980 | 0.1379 | 0.1724 | 0.2255 | 0.2346 |
| BB | 0.3209 | 0.2968 | | | | | | | |
| M3a | −0.2450 | −0.2355 | | | | | | | |
| M3b | 0.1736 | 0.1661 | 0.2831 | | | | | | |
| M4a | | 0.0622 | | | | | | | |
| M4b | 0.4658 | 0.4492 | 0.5654 | | | | | | |
| M6 | | | | | 0.0066 | 0.1460 | 0.1527 | 0.0543 | |
| DB | | | | | | | | | 0.0977 |
| M8a | | | | 0.3654 | | | | | |
| M9a | | | | 0.4289 | | | | | |
| Constant | −36.965 | −37.485 | −29.236 | −21.640 | −21.822 | −19.164 | −18.924 | −18.760 | −21.359 |
| Sectioning point | −0.0632 | −0.0632 | −0.0632 | −0.7499 | 0 | 0 | 0 | 0 | 0 |
| Accuracy M | 84.8 | 84.8 | 82.6 | 80.4 | 95.2 | 80.4 | 89.4 | 84.1 | 85.0 |
| F | 83.7 | 83.7 | 87.8 | 77.6 | 92.0 | 91.4 | 85.9 | 87.6 | 84.6 |

*Pop* population, *M* male, *F* female

Evolution (GAME) [58, 59]. The GAME method was utilized to search for new classification functions. This inductive method is based on a feed-forward artificial neural network and consists of different types of transfer functions. Both, the structure of the network and the parameters of the transfer functions, are set automatically during a run of GAME. The final structure of the network represents the new classification function. An example of such a network structure is shown in figure (Fig. 2). Because of the random initialization, the GAME method provides different discriminant functions in each run of the method. The result DF can be either linear or nonlinear. To obtain representative results from the GAME method, we used ten times repeated tenfold cross-validation.

Apart from the classification functions, the GAME method encapsulates also a feature ranking method called FeRaNGA [60]. The FeRaNGA method provides

**Table 4** Descriptive statistics of the measured variables of the tibia from Czech population

| | Male—descriptive statistics, $N = 30$ | | | | Female—descriptive statistics, $N = 26$ | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Dimension (mm) | Mean | SD | Min | Max | Mean | SD | Min | Max | *t* test |
| M3 | 81.67 | 4.01 | 72.51 | 90.02 | 76.85 | 5.29 | 69.92 | 92.10 | 0.000 |
| BB | 77.72 | 3.99 | 69.48 | 85.72 | 72.47 | 4.75 | 64.57 | 84.20 | 0.000 |
| M3a | 34.13 | 2.21 | 30.15 | 39.71 | 31.91 | 2.39 | 27.66 | 36.49 | 0.001 |
| M3b | 34.04 | 2.91 | 29.05 | 40.43 | 31.79 | 2.41 | 28.43 | 39.20 | 0.003 |
| M4a | 49.40 | 3.24 | 41.44 | 55.32 | 45.89 | 3.74 | 40.69 | 56.22 | 0.001 |
| M4b | 39.57 | 3.43 | 31.69 | 45.61 | 36.42 | 2.61 | 32.13 | 42.11 | 0.000 |
| M6 | 54.93 | 3.52 | 48.91 | 63.74 | 50.43 | 2.56 | 43.16 | 56.91 | 0.000 |
| DB | 50.04 | 2.93 | 45.25 | 55.59 | 48.29 | 3.50 | 41.42 | 56.40 | 0.050 |
| M8a | 37.74 | 3.95 | 30.20 | 46.37 | 34.91 | 3.19 | 29.69 | 43.22 | 0.006 |
| M9a | 28.32 | 2.18 | 24.91 | 33.32 | 26.09 | 2.29 | 23.19 | 32.62 | 0.001 |

*N* number of individuals, *SD* standard deviation, *Min* minimum, *Max* maximum

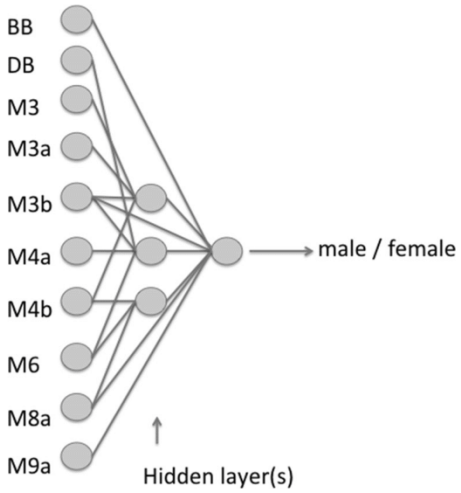


**Fig. 2** Example structure of the GAME neural network with one input layer, one hidden layer, and one output layer for classification of males and females. The number of hidden layers can differ for each particular run of the algorithm

importance of particular input variables used during the ten times repeated tenfold cross-validation process while searching for classification functions.

## Results

### Discriminant analyses

*Computation of DFA in Czech sample*

We designed two linear discriminant function equations (Table 5) for modern Czech population. First of them (CF1) includes seven variables (BB, M3, M3a, M3b, M4a, M4b, and M6) and provides 85.7 % (83.3 % males, 88.5 % females) of correct determination (sex bias −5.2 %). The second uses five variables (BB, M3a, M3b, M4a, and M4b) and determines sex with 82.1 % (80.0 % males, 84.6 % females) of correct assessment (sex bias −4.6 %). Cross-validation is not identical to the original data; in the first case, value is equal to 75.0 % and in the second to 76.8 %. The sectioning point is equal to 0, decision value above 0 indicate a male, while values below zero indicate a female.

*Advanced classification technique (GAME) in Czech sample*

The GAME method creates automatically a structure of the artificial neural network. The structure is represented by a classification function. Due to a random initialisation of GAME at the beginning, the classification function is different for each run of the method. From this reason, we used ten times repeated tenfold cross-validation (thus, we created 100 classification functions) to get mean classification accuracy of the GAME method for representative comparison with another approaches. Each particular classification function has different level of complexity. It can be either linear, nonlinear or any combination of both types. Classification function with the highest classification accuracy does not necessarily means that it is also applicable in practice (e.g., an exponential function of a logarithmic function having as an argument a sigmoidal function is not the right candidate for practice). The overall average classification accuracy from 100 runs of the GAME method was 82.35 % with a standard deviation of 2.25 %. Based on above mentioned we selected three user friendly and practically applicable classification functions with high classification accuracy while having minimum complexity. One linear classification function (CF3) and two nonlinear classification functions (CF4, CF5) are shown in Table 5. The linear CF3 includes six variables (DB, M3a, M4a, M4b, M6, M9a) and provides classification accuracy of 83.9 % (males 83.3 %, females 84.6 %) with sex bias of −1.3 %. The first of our nonlinear functions is CF4, and its type is exponential. The classification accuracy of CF4 is the same as for CF3 with the same sex bias, though it needs only four variables to reach such accuracy. The last classification function CF5 is of sigmoidal type and provides the highest classification accuracy (from CF1 to CF5) with 87.5 % (male 90 %, female 84.6 %) and sex bias 5.38 %.

The discriminatory value for CF3, CF4, and CF5 is equal to 0.5. Smaller value indicates male subject, while greater or equal values indicates female subject.

The importance of particular input variables determined by the FeRaNGA method during the search for best classification function are listed from the most important to the least important followed by their importance: M6 (28.2 %), DB (21.7 %), BB (16.4), M3 (14.9 %), M9a (8.7 %), M3a (4.8 %), M4a (2.5 %), M4b (1.8 %), M8a (0.6 %), and M3b (0.4 %). The four most important variables are M6, DB, BB, and M3.

*Application of dimensions from Czech sample into published DFA*

Table 6 compares classification success rates in the original reference sample and in target sample (Czech population) separately for males and females. When we used Czech dimensions in DF1, originally created for Portuguese sample, almost all of Czech males (90 %) were determined correctly, only three of them were by this function considered as female. On the other hand, females were overestimated and determined as male. Only 38.5 % of female sample was successfully classified. We got very similar results using DF2. Success of classification among males was 93.3 % and among females coincidently 38.5 %. DF3 classified correctly 83.3 % of males and more than half (53.9 %) of females. The last tested function originally proposed for Portuguese population,

**Table 5** The proposed classification functions in the Czech population

| | | Classification functions | | | | |
|---|---|---|---|---|---|---|
| | | CF1 | CF2 | CF3 | CF4 | CF5 |
| Variable/no. of used variables in CF | | 7 | 5 | 6 | 4 | 3 |
| M3 | | −0.434 | | | | * |
| BB | | 0.576 | 0.199 | | * | |
| M3a | | 0.005 | 0.102 | 0.004 | | |
| M3b | | −0.162 | −0.105 | | | |
| M4a | | 0.056 | 0.011 | −0.042 | | |
| M4b | | 0.169 | 0.163 | −0.042 | | * |
| M6 | | 0.395 | | −0.071 | * | |
| DB | | | | 0.073 | * | * |
| M9a | | | | −0.016 | * | |
| Constant | | −33.724 | −21.565 | 4.315 | | |
| Sectioning point | | 0 | 0 | 0.5 | 0.5 | 0.5 |
| Classification success (accuracy) | Male | 83.3 % | 80.0 % | – | – | – |
| | Female | 88.5 % | 84.6 % | – | – | – |
| | Average | 85.7 % | 82.1 % | – | – | – |
| Cross-validated | Male | – | – | 83.3 % | 83.3 % | 90.0 % |
| | Female | – | – | 84.6 % | 84.6 % | 84.6 % |
| | Average | 75.0 % | 76.8 % | 83.9 % | 83.9 % | 87.5 % |
| Sex bias | | −5.2 % | −4.6 % | −1.3 % | −1.3 % | −5.4 % |

CF4 is exponential function and CF5 is sigmoidal, for that reason they are presented out off table directly beneath. Used variables in those cases are marked by "*"

$$CF4 = 2.058 * \mathrm{EXP}(-0.041 * M6 - 0.052 * M9a + 3.516) - 0.017 * BB + 0.065 * DB - 3.452$$

$$CF5 = \frac{1}{1+\exp(-23.18*DB+70.676*M3+3.642*M4b-4489.232)}$$

DF4, estimated males flawlessly (100 %) but totally overestimated females of our sample (0 %). DF5 was fitted to Spanish population and separated properly 100 % of males and just 11.5 % of females. Discriminant function 6 derived from Italian population was not eventually incorporated (see below). DF7 was proposed on Greek population, and this function estimated absolutely correctly in our male sample, and largely overestimated female sample (3.9 %). DF8 was derived from mixed south European populations, and this function gave 100 % of correct assessment of males and 7.7 % of correct determination of females. The last tested discriminant function DF9 proposed for Euro-American population successfully classified all Czech males but only 7.7 % in female sample.

**Table 6** Application of DFA proposed in different populations in recent Czech population: simulation of disregarding DFA population specificity

| | Classification success in reference sample (%) | | | | Classification success in Czech sample (%) | | | |
|---|---|---|---|---|---|---|---|---|
| Discriminant function | Male | Female | Pooled mean | Sex bias | Male | Female | Pooled mean | Sex bias |
| DF1 (Portuguese) | 84.8 | 83.7 | 84.2 | 1.1 | 90.0 | 38.5 | 66.1 | 51.5 |
| DF2 (Portuguese) | 84.8 | 83.7 | 84.2 | 1.1 | 93.3 | 38.5 | 67.9 | 54.8 |
| DF3 (Portuguese) | 82.6 | 87.8 | 85.3 | −5.2 | 83.3 | 53.9 | 69.7 | 29.4 |
| DF4 (Portuguese) | 80.4 | 77.6 | 79.0 | 2.8 | 100.0 | 0.0 | 53.6 | 100 |
| DF5 (Spanish) | 95.2 | 92.0 | 93.4 | 3.2 | 100.0 | 11.5 | 58.9 | 88.5 |
| DF7 (Greek) | 89.4 | 85.9 | 87.8 | 3.5 | 100.0 | 3.9 | 55.4 | 96.1 |
| DF8 (pooled) | 84.1 | 87.6 | 86.0 | −3.5 | 100.0 | 7.7 | 57.1 | 92.3 |
| DF9 (Euro-American) | 85.0 | 84.6 | 84.8 | 0.4 | 100.0 | 7.7 | 57.1 | 92.3 |

The same table compares computed sex biases of reference and target sample. While values of published DFA are low (<7.5 %), sex biases of tested functions are unacceptably high (above 29.4 %).

## Discussion

Forensic anthropology benefits from accurate and reliable methods for estimating parameters of an individual's biological profile. The Daubert standard requires forensic anthropologists to confirm the validity of their methodologies through empirical testing, peer review, publication, and calculating error rates. External validity must be tested, and if the proposed procedures are based on one population, application to other population is required [1, 61, 62]. In cases of natural mass disasters, accidents, and mutilated human remains, sex estimation is often based on single skeletal element, most frequently on the long bones of extremities, vertebrae, or skull fragments [63, 64]. The success of determining sex is dependent not only on the method and degree of sexual dimorphism of the reference population in which the method was proposed, but also on the degree of sexual dimorphism of the target population in which the method is applied. Moreover the reliability of the sex allocation is influenced by secular trend [27, 28]. However, for these reasons, as we demonstrated in the Introduction, most of the methods for sex estimation are population-specific. At the same time, the tibia may be more susceptible to short-time secular changes and thus, more research is required to test this hypothesis [21].These short-term changes in dimensions of the tibia are also affected by the increasing prevalence of obesity [65]. A number of methods have been proposed on the identified skeleton collections that come from populations that do not longer exist. A number of such studies were published on historical skeletal collection and their methods are offered and available to users [8–15, 66]. As noted by Ousley and Jantz (2012), "Data from modern American show that standards derived from the nineteenth century are not appropriate for the assessment of twentieth-century groups." and "Also, postcranial samples from the Hamann-Todd Collection used by Iscan and Cotton (1990) perform very poorly when applied to modern cases" [67]. However, a major problem with use of the sex classification tools is that the population is not stable and is changed continuously under the influence of modifications in socioeconomic factors. Therefore, it is important to know the risk of errors, when the population specificity of the methods is completely ignored.

Altogether, we proposed five classification functions in Czech sample to determine sex (two discriminant functions, one linear, exponential, and sigmoidal). Maximum classification accuracy using a linear DF analysis was 85.7 % and −5.2 % sex bias with seven variables included (CF1). However, the highest classification accuracy of 87.5 % with −5.4 % sex bias and comprising three variables (CF5) was achieved using feed-forward artificial neural network–GAME method. This confirms that other classification functions than classic linear DFA are suitable as classification methods [24, 68]. However, with all five functions, we reach an acceptable sex bias according to Franklin et al. [25], ranging from −1.3 to −5.4 %, only two of them slightly exceed the 5 % limit. More importantly, we have achieved comparable success rate and sex bias (Table 5) as other mentioned authors did in other population samples. Similar results for samples from twentieth and twenty-first centuries reached also Brzobohatá et al. [52] using geometric morphometry. Although CF5 (sigmoidal function) provides the best determination of sex, it has the greatest sex bias (5.4 %), which is still acceptable. It should be noticed that CF3 (linear) and CF4 (exponential) have the same accuracy and even sex bias, but CF3 includes six variables and CF4 only four variables. Generally, nonlinear classification functions (CF4 and CF5) tend to require less input variables. This is advantageous, at least in cases of incomplete or damaged remains, where it is not possible to measure all dimensions.

In this study, we tested accuracy of discriminant functions fitted to various populations with dimensions measured on our Czech. We did it in the same manner as a potential user when estimating the sex from the tibia. We used only multivariable methods because a single variable has little practical value for sex assessment according to Peckmann et al. [37] due to the large overlapping area of both sexes. Results reveal strong sexual dimorphism of the tibial dimensions in the Czech population which significantly contributes to sex discrimination.

When classifying the Czech sample using seven DF from three different Mediterranean European populations [14, 21] or DF from American Caucasians [10], sex estimation was done correctly only in 53.6–69.7 % of cases. However, regardless of the rate of misclassification that is unacceptable for practical use in the forensic sciences, the sex bias ranged from 29.4 to 100 %. The results repeatedly demonstrated on the one hand, almost absolutely correct determination of males (DF1-5, DF7-9) and on the contrary, failure of classification of female individuals. DF6 is incorrect, and there is a mistake in coefficients or constant computed for function IF4 in the original paper of Kranioti and Apostol [21]. It can be verified by solving discriminant equation with mean values; in that case, Italians of both sexes have negative values despite sectioning point of zero.

The low classification rates in females were observed also in sexing of the metacarpals [69], and results in this study indicate that the standards developed from the continental Greece are not proper for application in forensic cases in Crete because they do not represent the local Cretan population. The original

sectioning point sometimes completely misses the distribution of the sexes. Low accuracy together with high sex bias is unacceptable for forensic anthropology. Similar results, proving that population differences cause that DF that perform well on one population may produce high error rates and sex biases when applied to other populations, were noted by several authors in crania [55, 70, 71].

The most likely explanations of such results are the variability of individual populations [72], different degree of sexual dimorphism of tibia among populations [73] and thus different sex-specific means of dimensions of each population (variance of values dimensions). We should take into consideration that unlike chosen studies, where DF1-DF9 were proposed, we used 3D models of bones. According to previous studies [74, 75], where authors support their comparability on skulls, we believe that measurements derived from 3D models are comparable to those from dry bones. Therefore, the general effect of applying nonpopulation standards reduces the accuracy of classification, the magnitude of which is proportionately related to the degree of divergence in size between the original reference sample and the individual to which those standards will be applied. It follows that the main limitation of using published methods for estimating sex using the metric data for tibia (and other bones) is their population specificity (e.g., [21, 42]). Evidence follows from contradiction of comparing the results of reliability and accuracy of classification methods. Population specificity is also related to the trend of globalization [76]. Problem is that the individual whose sex we want to estimate does not necessarily originally come from the area in which their remains were found (e.g., [4, 77]). The population affinity or ancestry of an unknown specimen in the worldwide increase in admixture of human populations is difficult (e.g., [78–80]).

What options are there for dealing with population specificity of methods which is characteristic not only of sex estimation methods but also of all methods estimating the biological profile of the individual? Forensic anthropology is constrained by a paucity of population-specific standards as number of repositories of documented skeletons, traditionally the main source of population-specific data, is limited [62]. However, as Albanese et al. [81] mentioned, group specificity should increase the precision of the estimation methods, but on the other hand, in many cases, the parameters of a given group are based on assumptions which make assigning an unknown to a group under the best conditions problematic and at worst impossible. Methods developed in pooled samples acquired, contrary to traditional assumptions, certain degree of robustness. For example, pooled-ancestry stature equation can be more appropriate than population-specific equations obtained from a single sample [82]. It should be kept in mind that it is not possible to respect the population specificity in bioarchaeology. The only possible solution could be the creation of regional standards in Europe, which correspond to variability of body size, and on such a basis, a robust method for sex estimation should be developed. It was similarly designed in the study concerning stature in bioarchaeology [83], where the authors created northern and southern European formulae.

Medical imaging (CT, MRI, ultrasonography) of living individuals offer appropriate and reliable source of contemporary population data from selected geographic populations. These data can be pooled in a metapopulation sample from which skeletal standards for the estimation of age, sex, and stature could be developed. The combination of recent data from the CT images and historical collections of identified individuals could lead to suggestions of highly robust methods, whose reliability would provide a reliable estimation of the biological profile in forensic anthropology and bioarchaeology. For these reasons, one of the goals of the international scientific community is to develop a synergistic, well-coordinated activity to create a database of 3D imaging from clinical computed tomographic scans of the skeleton to formulate new nonpopulation-specific methods.

## Conclusion

Sexual dimorphism of the tibia is well established in the literature and standards for the Czech have been proposed. However, all the DFA proposed in Czech sample and all of already published functions in other populations were population-specific with classification accuracy ranging between 80 and 90 %. Disregarding or ignoring this fact causes complete failure of sex estimation. Direct application of published population-specific DFA in Czech sample leads to unacceptably low accuracy with sex bias ranging from −29.4 to 100 %. This failure of sex assessment is caused by a variability of skeletal dimensions among populations. For these reasons, the misclassification rate does not correspond to the Daubert criteria, and therefore, the DFA from tibia cannot be recommended also for the estimation of sex of recent skeletal remains in Europe.

**Acknowledgments** We would like to thank anonymous reviewer since he revealed that in the original article of Kranioti and Apostol must be numerical mistake in one proposed discriminant function.

## References

1. Cattaneo C (2007) Forensic anthropology: developments of a classical discipline in the new millenium. Forensic Sci Int 165:185–193
2. Dawson C, Ross D, Mallet X (2011) Sex determination. In: Black S, Ferguson E (eds) Forensic Anthropology: 2000 to 2010. CRC Press, Boca Raton, pp 61–94

3. Moore MK (2013) Sex Estimation and Assessment. In: DiGangi EA, Moore MK (eds) Research Methods in Human Skeletal Biology. Academic Press, London, pp 91–116
4. Spradley MK, Jantz RL, Robinson A, Peccerelli F (2008) Demographic Change and Forensic Identification: Problems in Metric Identification of Hispanic Skeletons. J Forensic Sci 53:21–28
5. Özer BK, Özer I, Sağır M, Güleç E (2014) Sex determination using the tibia in an ancient anatolian population. Med Archaeol Archaeom 14:329–336
6. Novotný V, Işcan MY, Loth SR (1993) Morphologic and osteometric assesment of age, sex and race from the skull. In: Işcan MY, Helmer RP (eds) Forensic analyses of the skull. Wiley–Liss Inc., New York, pp 71–88
7. Dangar KP, Pandya AM, Rathod SP, Tank KC, Akbari VJ, Solanki SV (2012) Sexual dimorphism of proximal epiphyseal breadth of tibia. Int J Biol Med Res 3:1331–1334
8. Işcan MY, Miller-Shaivitz P (1984a) Determination of sex from the tibia. Am J Phys Anthropol 64:53–57
9. Albanese J, Eklics G, Tuck A (2008) A metric method for sex determination using the proximal femur and fragmentary hipbone. J Forensic Sci 53:1283–1288
10. Işcan MY, Miller-Shaivitz P (1984b) Discriminant function sexing of the tibia. J Forensic Sci 29:1087–1093
11. Holland TD (1991) Sex assessment using the proximal tibia. Am J Phys Anthropol 85:221–227
12. Kieser JA, Moggi-Cecchi J, Groeneveld HT (1992) Sex allocation of skeletal material by analysis of the proximal tibia. Forensic Sci Int 56:29–36
13. Işcan MY, Yoshino M, Kato S (1994) Sex determination from the tibia: standards from contemporary Japan. J Forensic Sci 39:785–792
14. Brůžek J (1995) Diagnose sexuelle à l'aide de l'analyse discriminante appliquée au tibia. Antropol Port 13:93–106
15. Steyn M, Işcan MY (1997) Sex determination from the femur and tibia in South African whites. Forensic Sci Int 90:111–119
16. Gonzales-Reimers E, Velasco-Vázquez J, Arnay-de-la- Rosa M, Santolaria Fernandez F (2000) Sex determination by discriminant function analysis of the right tibia in the prehistoric population of the Canary Islands. Forensic Sci Int 108:165–172
17. Sakaue K (2004) Sexual determination of long bones in recent Japanese. Anthropol Sci 112:75–81
18. Šlaus M, Tomičić Ž (2005) Discriminant function sexing of fragmentary and complete tibiae from medieval Croatian sites. Forensic Sci Int 147:147–152
19. Garcia S (2010) Is the circumference at the nutrient foramen of the tibia of value to sex determination on human osteological collections? Testing a new method. Int J Osteoarchaeol 22:361–365
20. Šlaus M, Bedić Ž, Strinović D, Petrovečki V (2013) Sex determination by discriminant function analysis of the tibia for contemporary Croats. Forensic Sci Int 226:302e1–e4
21. Kranioti EF, Apostol MA (2015) Sexual dimorphism of the tibia in contemporary Greeks, Italians, and Spanish: forensic implications. Int J Legal Med 129:357–363
22. Iscan MY, Steyn M (2013) The Human Skeleton in Forensic Medicine. Charles C. Thomas Publisher, Springfield, p 493
23. Robinson MS, Bidmos MA (2011) An assessment of the accuracy of discriminant function equations for sex determination of the femur and tibia from a South African population. Forensic Sci Int 206:212.e1–212.e5
24. Santos F, Guyomarc'h P, Brůžek J (2014) Statistical sex determination from craniometrics: Comparison of Linear Discriminant Analysis, Logistic Regression, and Support Vector Machines. Forensic Sci Int 245:204.e1–204.e8
25. Franklin D, Cardini A, Flavel A, Kuliukas A (2013) Estimation of sex from cranial measurements in a Western Australian population. Forensci Sci Int 229:158.e1–158.e8
26. Buikstra JE, Ubelaker DH (1994) Standards for data collection from human skeletal remains. Arkansas Archaeological Survey Research, Fayetteville
27. Jantz RL (2001) Cranial change in Americans: 1850–1975. J Forensic Sci 46:784–787
28. Meadows Jantz L, Jantz RL (1999) Secular change in long bone length and proportion in the United States, 1800–1970. Am J Phys Anthropol 110:57–67
29. Meadows L, Jantz RL (1995) Allometric secular change in the long bones from the 1800s to the present. J Forensic Sci 40:762–767
30. Saini V (2014) Significance of temporal changes on sexual dimorphism of cranial measurements of Indian population. Forensic Sci Int 242:300e1–300e8
31. Mascie-Taylor CGN, Bogin B (1995) Human variability and plasticity. Cambridge University Press, Cambridge
32. Ishak NI, Hemy N, Franklin D (2012) Estimation of stature from hand and handprint dimensions in a Western Australian population. Forensic Sci Int 216:199. e1–199. e7
33. Danubio ME, Sanna E (2008) Secular changes in human biological variables in Western Countries: An updated review and synthesis. J Anthropol Sci 86:91–112
34. Sanna E, Milia N, Martella P, Danubio M (2015) Body and head dimensions of adults in Sardinia (Italy) support different intensities of relative secular trends. J Anthropol Sci 93:157–162
35. Bigoni L, Veleminska J, Brůžek J (2010) Three-dimensional geometric morphometric analysis of cranio-facial sexual dimorphism in a Central European sample of known sex. Homo 61:16–32
36. Timonov P, Fasová A, Radoniova D, Alexandrov A, Delev D (2014) A study of sexual dimorphism in the femur among contemporary Bulgarian population. Euras J Anthropol 5:46–53
37. Peckmann TR, Orr K, Meek S, Manolis SK (2015) Sex determination from the calcaneus in a 20th century Greek population using discriminant function analysis. Sci Justice 55:377–382
38. Giles E, Elliot O (1963) Sex determination by discriminant function analysis of crania. Am J Phys Anthropol 21:53–68
39. Steyn M, Iscan MY (1998) Sexual dimorphism in the crania and mandibles of South African whites. Forensic Sci Int 98:9–16
40. Franklin D, Flavel A, Kuliukas A, Oxnard CE, Marks MK (2012) Cranial sexual dimorphism and anthropological standards: Preliminary investigations in a Western Australian population. 64th Annual Meeting of the American Academy of Forensic Sciences, Atlanta
41. Králík M, Urbanová P, Wagenknechtová M (2014) Sex assessment using clavicle measurements: Inter- and intra-population comparisons. Forensic Sci Int 234:181e1–181e15
42. Bidmos MA, Dayal MR (2004) Further evidence to show population specificity of discriminant function equations for sex determination using the talus of South African blacks. J Forensic Sci 49: 1165–1170
43. Ross AH, Ubelaker DH, Kimmerle EH (2011) Implications of dimorphism, population variation, and secular change in estimating population affinity in the Iberian Peninsula. Forensic Sci Int 206: 214 e1–214 e5
44. Steyn M, Patriquin ML (2009) Osteometric sex determination from the pelvis - Does population specificity matter? Forensic Sci Int 191:113e1–113e5
45. Guyomarc´h P, Velemínská J, Sedlak P, Dobisíková M, Švenkrtová I, Brůžek J (2016) Impact of secular trend on sex assessment evaluated through femoral dimensions in a Czech population. Forensic Sci Int Sci Int. doi:10.1016/j.forsciint.2016.02.042
46. Dedouit F, Telmon N, Costagliola R, Otal P, Florence LL, Joffre F, Rougé D (2007) New identification possibilities with postmortem multislice computed tomography. Int J Legal Med 121:507–510
47. Ramsthaler F, Kettner M, Gehl A, Verhoff MA (2010) Digital forensic osteology: Morphological sexing of skeletal remains using volume-rendered cranial CT scans. Forensic Sci Int 195:148–152

48. Stull KE, Tise ML, Ali Z, Fowler DR (2014) Accuracy and reliability of measurements obtained from computed tomography 3D volume rendered images. Forensic Sci Int 238:133–140
49. Gulhan O, Harrison K, Kiris A (2015) A new computer-tomography-based method of sex estimation: Development of Turkish population-specific standards. Forensic Sci Int 255:2–8
50. Mesteková Š, Brůžek J, Veleminská J, Chaumoitre K (2015) A Test of the DSP Sexing Method on CT Images from a Modern French Sample. J Forensic Sci 60:1295–1299
51. Clavero A, Salicrú M, Turbón D (2015) Sex prediction from the femur and hip bone using a sample of CT images from a Spanish population. Int J Legal Med 129:373–383
52. Brzobohatá H, Krajíček V, Horák Z, Velemínská J (2015) Sex Classification Using the Three-Dimensional Tibia Form or Shape Including Population Specificity Approach. J Forensic Sci 60:29–40
53. Martin R, Saller K (1957) Lehrbuch der Anthropologie in systematischen darstellung mit besonderer berücksichtigung der anthropologischen methoden. Band I, Gustav Fischer Verlag, Stuttgart
54. Sjøvold T (1988) Geschlechtsdiagnose am Skelett. In: Martin R, Knussmann R (eds) Anthropologie, Handbuch des vergleichenden Biologie des Menschen. Gustav Fischer Verlag, Stuttgart, pp 444–480
55. Walker PL (2008) Sexing skulls using discriminant function analysis of visually assessed traits. Am J Phys Anthropol 136:39–50
56. Franklin D, Cardini A, Flavel A, Kuliukas A, Marks MK, Oxnard CE, O'Higgins P (2012) Estimation of sex from sternal measurements in a Western Australian population. Forensic Sci Int 217: 230.e1–230.e5
57. Statistica [software]. version 12 (www.statsoft.cz).
58. Kordík P (2009) GAME–hybrid self-organizing modeling system based on GMDH. In: Onwubolu GC (ed) Hybrid Self-Organizing Modeling Systems. Springer-Verlag, Heidelberg, pp 233–280
59. Kordík P, Koutník J, Drchal J, Kovářík O, Čepek M, Šnorek M (2010) Meta-learning approach to neural network optimization. Neural Netw 23:568–582
60. Pilný A, Kordík P, Šnorek M (2008) Feature Ranking Derived from Data Mining Process. In: Proceedings of the 18th international conference on Artificial Neural Networks, Part II. Springer-Verlag, Heidelberg, pp 889–898
61. Grivas CR, Komar DA (2008) Kumho, Daubert, and the Nature of Scientific Inquiry: Implications for Forensic Anthropology. J Forensic Sci 53:771–776
62. Franklin D, Cardini A, Flavel A, Kuliukas A (2012) The application of traditional and geometric morphometric analyses for forensic quantification of sexual dimorphism: preliminary investigations in a Western Australian population. Int J Legal Med 126:549–558
63. Gama I, Navega D, Cunha E (2015) Sex estimation using the second cervical vertebra: a morphometric analysis in a documented Portuguese skeletal sample. Int J Legal Med 129:365–372
64. Masotti S, Succi-Leonelli E, Gualdi-Russo E (2013) Cremated human remains: is measurement of the lateral angle of the meatus acusticus internus a reliable method of sex determination? Int J Legal Med 127:1039–1044
65. Harrington KI, Wescott DJ (2015) Size and Shape Differences in the Distal Femur and Proximal Tibia between Normal Weight and Obese American Whites. J Forensic Sci 60:S32–S38
66. Albanese J (2013) A method for estimating sex using the clavicle, humerus, radius, and ulna. J Forensic Sci 58:1413–1419
67. Ousley D, Jantz RL (2012) Fordisc 3 and statistical methods for estimating sex and ancestry. In: Dirkmaat DC (ed) A Companion to Forensic Anthropology. Wiley-Blackwell, Malden
68. Navega D, Vicente R, Vieira DN, Ross AH, Cunha E (2015) Sex estimation from the tarsal bones in a Portuguese sample: a machine learning approach. Int J Legal Med 129:651–659
69. Nathena D, Gambaro L, Tzanakis N, Michalodimitrakis M, Kranioti EF (2015) Sexual dimorphism of the metacarpals in contemporary Cretans: Are there differences with mainland Greeks? Forensic Sci Int 257:515.e1–515.e2
70. Messer DL, Ousley SD, Tuamsuk P (2013) Pattern in worldwide craniometric sexual dimorphism and its importance in forensic anthropology. in: Proceedings of the American Academy of Forensic Sciences XIX, pp 394–395
71. Franklin D, Flavel A, Kuliukas A, Hart R, Marks MK (2012) The development of forensic anthropological standards in Western Australia. in: Proceedings of the American Academy of Forensic Sciences XVIII, pp 361–362
72. White TD, Folkens PA (2005) The human bone manual. Academic Press, San Diego
73. Rösing FW, Graw M, Marré B, Ritz-Timme S, Rotschild MA, Rotzscher K, Schmeling A, Schroder I, Geserick G (2007) Recommendations for the forensic diagnosis of sex and age from skeletons. Homo 58:75–89
74. Citardi MJ, Herrmann B, Hollenbeak CS, Stack BC, Cooper M, Bucholz RD (2001) Comparison of scientific calipers and computer-enabled CT review for the measurement of skull base and craniomaxillofacial dimensions. Skull Base 11:5–11
75. Verhoff MA, Ramsthaler F, Krähahn J, Deml U, Gille RJ, Grabherr S, Thali MJ, Kreutz K (2008) Digital forensic osteology–possibilities in cooperation with the Virtopsy project. Forensic Sci Int 174: 152–156
76. L´Abbé EN, Steyn M (2012) The Establishment and Advancement of Forensic Anthropology South Africa. In: Dirkmaat DC (ed) A Companion to Forensic Anthropology. Wiley-Blackwell, Malden, pp 626–638
77. Brůžek J, Murail P (2006) Methodology and realiability of sex determination from the skeleton. In: Schmitt A, Cunha E, Pinheiro J (eds) Forensic Anthropology and Medicine: Complementary Sciences From Recovery to Cause of Death. Humana Press Inc., Totowa, pp 225–242
78. Ramsthaler F, Kreutz K, Verhoff MA (2007) Accuracy of metric sex analysis of skeletal remains using Fordisc® based on a recent skull collection. Int J Legal Med 121:477–482
79. Urbanová P, Ross AH, Jurda M, Nogueira MI (2014) Testing the reliability of software tools in sex and ancestry estimation in a multi-ancestral Brazilian sample. Leg Med 16:264–273
80. Sierp I, Henneberg M (2015) Can ancestry be consistently determined from the skeleton? Anthropolo Rev 78:21–31
81. Albanese J, Tuck A, Gomes J, Cardoso HFV (2016) An alternative approach for estimating stature from long bones that is not population- or group-specific. Forensic Sci Int 259:59–68
82. Hatza AN, Ousley SD, Cabo LL (2015) Estimating stature when ancestry is unknown: What statistical methods work best? In: Proceedings of the American Academy of Forensic Sciences XXI, pp 152–153
83. Ruff CB, Holt BM, Niskanen M, Sladek V, Berner M, Garofalo E, Garvin HM, Hora M, Maijanen H, Niinimaki S, Salo K, Schuplerova E, Tompkins D (2012) Stature and body massestimation from skeletal remains in the European Holocene. Am J Phys Anthropol 148:601–617

# Appendix A.2

Forensic Science International 287 (2018) 163–175

Contents lists available at ScienceDirect

Forensic Science International

journal homepage: www.elsevier.com/locate/forsciint

# Age estimation of adult human remains from hip bones using advanced methods

Anežka Kotěrová[a,*], David Navega[b], Michal Štepanovský[c], Zdeněk Buk[c], Jaroslav Brůžek[a], Eugénia Cunha[b]

[a] Department of Anthropology and Human Genetics, Faculty of Science, Charles University, Viničná 7, Prague 128 43, Czech Republic
[b] Centre for Functional Ecology, Laboratory of Forensic Anthropology, Department of Life Sciences, University of Coimbra, Calçada Martim de Freitas, 3000-456 Coimbra, Portugal
[c] Faculty of Information Technology, Technical University in Prague, Thákurova 9, Prague 160 00, Czech Republic

ARTICLE INFO

*Article history:*
Received 15 January 2018
Received in revised form 20 March 2018
Accepted 28 March 2018
Available online 4 April 2018

*Keywords:*
Age-at-death estimation
Pelvis
Advanced mathematical methods
Pubic symphysis
Auricular surface
Visual assessment

ABSTRACT

The assessment of age-at-death is an important and challenging part of investigations of human skeletal remains. The main objective of the present study was to apply different mathematical approaches in order to reach more accurate and reliable results in age estimation. A multi-ethnic dataset (n = 941) of evaluated age-related changes on the pubic symphysis and the auricular surface of the hip bone was used. Two research groups examined nine different mathematical approaches. The best results were reached by Multi-linear regression, followed by the Collapsed regression model, with MAE values of 9.7 and 9.9 years, respectively, and with RMSE values of 12.1 and 12.2, respectively. The mean accuracy of decision tree models ranged between 30.7% and 72.3%, with the model using only the PUSx indicator performing the best. Moreover, our results indicate that the limiting factor of age estimation can be the visual evaluation of age-related changes. Further research is required to objectify the proposed methods for estimating age.

© 2018 Elsevier B.V. All rights reserved.

## 1. Introduction

One of the priorities of both forensic anthropology and bioarchaeology when dealing with unknown human remains is to create each individual's biological profile [1]. Besides sex, stature and ancestry, assessment of age-at-death is also mandatory. Whereas the aging of juveniles is not difficult to assess, because it relies on the development and growth of immature bones, which are predictable and therefore allow to estimate age with low bias and within an acceptable range of error [2,3]. For adults, however, this is not the case [4]. The estimation of age-at-death for adults is much more challenging because most methods for age determination are based on age-related changes such as skeletal remodeling or degenerative changes and dental wear [5,6], whose rates are variable between populations and even between different individuals [2,7,8]. The reasons for this reside in exogenous and endogenous factors such as genetics, the environment, activity, health [2,9,10] and even in body size [11–13]. Above all, the skeletal aging of adults is still insufficiently described and, implicitly, unpredictable.

In paleodemography, we are experiencing a slightly different perspective of age-at-death estimation, since there is a major effort to reconstruct the demographic behavior of past populations (e.g. [14,15]). In forensic anthropology, on the other hand, it is a priority to identify the skeletal remains of an individual, and forensic experts therefore estimate an individual's age [16,17] with a certain accuracy and reliability.

In general, the following skeletal age indicators are used: ribs [18–21], the clavicle [22] and the sacrum [23], the cranial suture closure [24] being the most frequently used articular surfaces of the hip bone. Primarily it is the pubic symphysis [9,25–28], where the Brooks and Suchey method is the most used one [29], and the auricular surface [7,28,30,31]. More recently, the acetabulum is at the forefront of research interest [32–36]. In addition, teeth are also used to estimate age-at-death of adult individuals. A special case is cementochronology, which evaluates the alternation of mineralized layers in dental cementum [37,38].

Among the innumerable age estimation methods, we can distinguish between single-indicator methods (e.g. [11–13,18–21,24]) and multiple-indicators methods [39–42], which are often recommended [43]. Different methods evaluate age-related changes according to a phase-based [18,25,27] or a component (composite)-based [7,31,44] scoring system. These two systems differ from each

* Corresponding author.
*E-mail address:* koterova.a@seznam.cz (A. Kotěrová).

https://doi.org/10.1016/j.forsciint.2018.03.047
0379-0738/© 2018 Elsevier B.V. All rights reserved.

other in the manner in which this information about morphological traits is translated into a useable format for age estimation. Component-based systems, which according to the results of Shirley and Ramirez Montes [45] provide more objective scoring, as they divide age indicator into different characteristics and score each of them separately. By contrast, phase-based systems describe age indicators as a whole and delimit broad phases. Nevertheless, the majority of methods are population-specific and largely derived from only a single population [27,31,32,46]. Moreover, most age estimation methods are based on visual assessment of age indicators, which is inevitably subjective. There are only a few exceptions that attempt to use mathematical approaches for the evaluation of age changes [47–49].

It can be said that we are contributing to an already excessively large array of methods for estimating the age of adults, each of which competes with the others by attempting to achieve better accuracy and precision [5]. Regardless of the number of already available methods, the problem of aging of older individuals remains, in particular those over 40–50 years [1,9,11], which is why broad intervals (40+, 50+, 60+ years) are often proposed [9,50], with an implicit loss of precision of age estimation techniques, precluding identification.

Recently, considerable attention has been paid to non-parametric methods and to new classification techniques such as decision trees, random forests and neural networks. Their use in anthropology could be highly beneficial as it would improve both accuracy and reliability. These procedures are already being used for other components of the biological profile [51–55], but are yet to be used extensively in age estimation. Exceptions are Dudzik and Langley [56] or Buk et al. [51] who proposed advanced data mining methods and found the sex of individuals not to be important for age estimation, at least in their dataset (which is the same as the one in the present study). Their results also indicate that accurate and reliable age estimation is possible only between three wide age classes (less than 30 years, 30–60 and over 60).

In the present study, two independent research groups took different mathematical approaches on appraised multi-population data to estimate age-at-death. Our goal is to reach more accurate classification and to divide the human life span into more than the three age intervals proposed by Buk et al. [51].

## 2. Material

A multi-ethnic database composed of nine samples (n = 941) from all over the world, with age ranging from 19 to 100 years, was used. It has to be emphasized that all the skeletons were identified. The population samples used have the following origins: Portugal, Switzerland, Spain, America (individuals of both European and African descent), Africa (Afrikaner, Zulu and Soto populations) and Thailand. The numbers of individuals in each sample are presented in Table 1. The Portuguese collection, known as the Coimbra identified Skeletal Collection CISC, is stored at the University in Coimbra and consists of individuals that lived in the 19th and 20th centuries [57]. The Simon collection from Switzerland is made of individuals that lived in the second half of the 20th century and is stored at the University of Geneva [58]. The Spanish collection comprises individuals that died during the 20th century and is stored at the Institute of Forensic Medicine, Complutense University of Madrid [59]. The Euro-American and Afro-American samples come from the Hamman-Todd collection, which is housed by the Cleveland Museum of Natural History in Ohio and includes individuals who lived in 19th–20th century [60]. African samples (Zulu, Soto and Afrikaners of European descent) originated from the Dart collection, which is deposited at the Department of Anatomy, the University of the Witwatersrand in Johannesburg. The individuals lived in the 19th century and during the first half of the 20th century [61]. The last sample, a Thai population from the late 20th century, is stored at the Department of Anatomy, University of Chiang Mai. The data were kindly provided by Aurore Schmitt.

**Table 1**
Osteological samples used in this study.

| Sample | Ancestry | Origin | Collection | Females | Males | Total |
|---|---|---|---|---|---|---|
| Portugal | Caucasian | Europe | Coimbra | 85 | 74 | 159 |
| Switzerland | Caucasian | Europe | Geneva | 17 | 26 | 43 |
| Spain | Caucasian | Europe | Madrid | 34 | 33 | 67 |
| USA 1 | Caucasian | America | Cleveland | 73 | 69 | 142 |
| USA 2 | Afroamerican | America | Cleveland | 40 | 41 | 81 |
| Afrikaner | Caucasian | South Africa | Johannesburg | 32 | 35 | 67 |
| Zulu | African | South Africa | Johannesburg | 102 | 102 | 204 |
| Soto | African | South Africa | Johannesburg | 44 | 33 | 77 |
| Thailand | Asian | Asia | Chang-Mai | 45 | 56 | 101 |
| | | | Total: | 472 | 469 | 941 |

## 3. Methods and age estimation models used

### 3.1. Scoring system of senescence skeletal changes

Two age indicators traditionally used for the estimation of age-at-death were chosen: the pubic symphysis and the sacro-pelvic surface. The scoring system used was developed by Schmitt and first published in her dissertation thesis [59] and was afterwards used in published articles [28,62]. Each feature of both indicators is processed and evaluated separately. A detailed description is presented in Table 2. All individuals were evaluated by one observer (Aurore Schmitt).

### 3.2. Collapsed regression model

Here we introduce a Collapsed regression model, which collapses all variables into one cumulative variable. Age is estimated according to the following exponential equation:

$$Age_{Estimated} = a_1 * exp(a_2 * CUMM),$$

where $a_1$ and $a_2$ are coefficients to be determined and CUMM is defined as the sum of all individual scores as follows:

$$CUMM = PUSA + PUSB + PUSC + SSPIA + SSPIB + SSPIC + SSPID,$$

where *PUSA*,*PUSB*, *PUSC*, *SSPIA*, *SSPIB* and *SSPID* represent numerical values of the scores as given in Table 2. Figs. 1 and 2 explain the relationship between age and CUMM variable. Fig. 1 shows age growing linearly while a logarithmic scale was used in case of Fig. 2.

### 3.3. Multi-linear regression

The Multi-linear regression model estimates the age of an individual based on the scoring of all variables according to the following equation:

$$Age = b_1 * PUSA + b_2 * PUSB + b_3 * PUSC + b_4 * SSPIA + b_5 * SSPIB + b_6 * SSPIC + b_7 * SSPID + b_8,$$

where constants $b_1, b_2, \ldots, b_8$ are computed using the least squares approach. Attributes that are collinear are removed. Attribute selection using the Akaike information metric is used to remove other attributes when acceptable. The relationship between age and variables PUSx and SSPIx is shown in Fig. 3. A fitted line for every variable approximates estimated age by linear equation. The correlation coefficient indicates a relationship between "score" and "actual age" for every variable.

**Table 2**
Summary table of scoring system for the pubic symphysis and the sacro-pelvic surface of the ilium.

| Variables | Scoring | | | |
|---|---|---|---|---|
| | Score 1 | Score 2 | Score 3 | Score 4 |
| Pubic symphysis (PUS) | | | | |
| Posterior plate (PUSA) | Ridges and Furrows | Surface flattening, appearance of dorsal margin | Complete | – |
| Ventral plate (PUSB) | Ridges and Furrows | Surface flattening, rampart expanding | Ventral rampart completely formed | – |
| Dorsal lip (PUSC) | Regular dorsal rim | Plain lipping | – | – |
| Surface sacro-pelvic (SSPI) | | | | |
| Transverse organization (SSPIA) | Undulation or striae present | Undulation or striae absent | – | – |
| Texture and porosity (SSPIB) | Dense surface | Uniform or partially coarsening granular texture | Coarse granular texture, porosities | Irregular surface, deep porosities, bone destruction |
| Apical activity (SSPIC) | Apex sharp, margins regular | Apex broad, rim formation | – | – |
| Retroauricular activity (SSPID) | Smooth surface | Irregular surface, osteophytes | – | – |

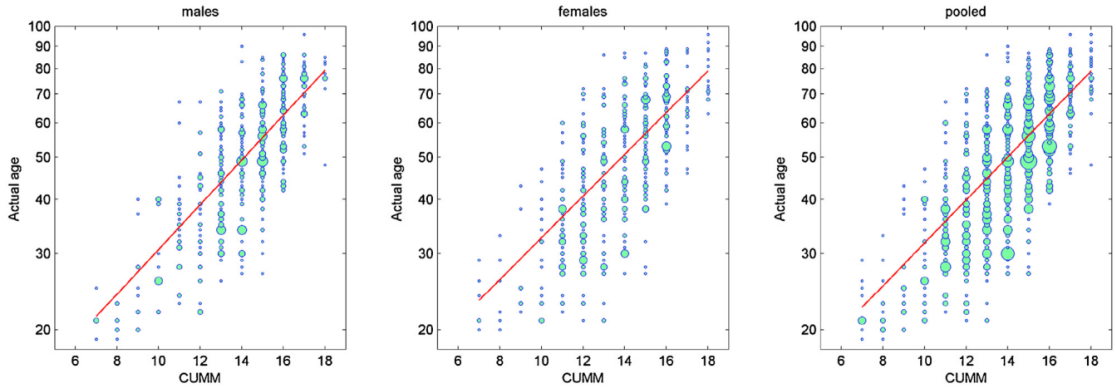


**Fig. 1.** Linear relationship between age and variable CUMM, for female, male and pooled dataset.
Size of the spot corresponds to the incidence of the pair — age and CUMM.

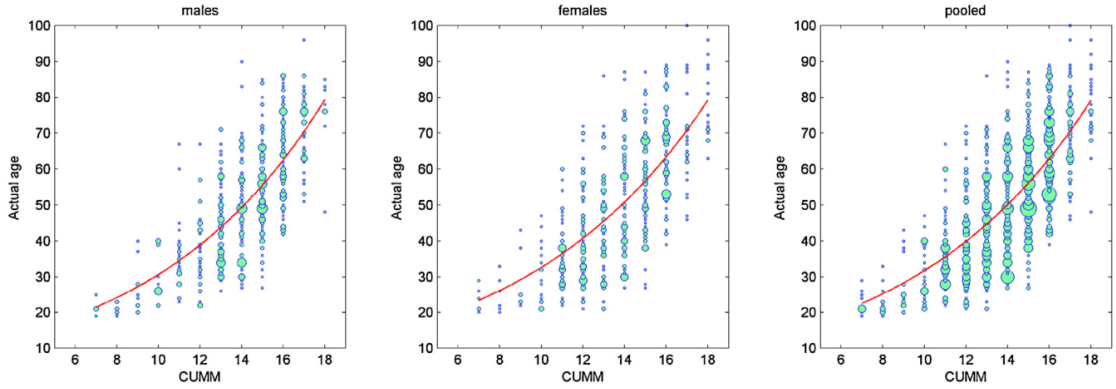


**Fig. 2.** Relationship between age and variable CUMM expressed by logarithmic scale.
Size of the spot corresponds to the incidence of the pair — age and CUMM.

### *3.4. Interval-based model*

In this section we introduce a novel interval-based model for age estimation. This interval-based model directly uses age intervals for each observed variable and scoring from the input dataset. The age interval covers less than 100% of all observed data in order to avoid anomalies if they exist. In this case, age intervals according to the sex of the person examined are used (similar to Table 3). The age interval is specified as <min, max> for each variable and score. If we define a function F such that the output of the function corresponds to the number of overlapping intervals, then we can estimate the age by 4th-order polynomial

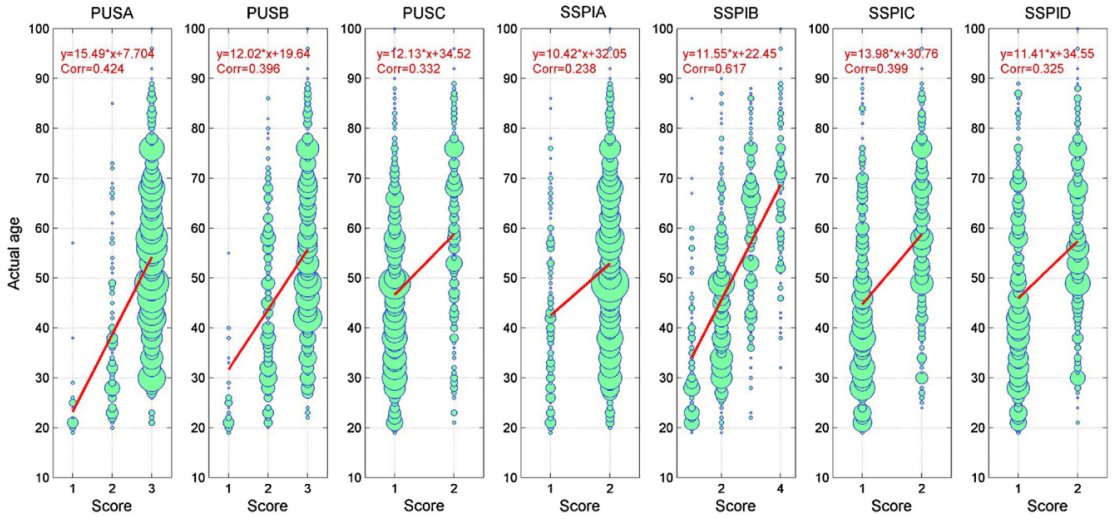


**Fig. 3.** Relationship between age and the explanatory variables (PUSx and SSPIx) with a fitted line, pooled dataset. Different diameter of spots correspond to the number of individuals of given age with specific score.

**Table 3**
Representative values for 80% coverage of the data for each variable and scoring (in years).

| Score | PUSA | | | PUSB | | | PUSC | | | SSPIA | | | SSPIB | | | SSPIC | | | SSPID | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | min | max | med | min | max | med | min | max | med | min | max | med | min | max | med | min | max | med | min | max | med |
| 1 | 20 | 29 | 21 | 20 | 34 | 23 | 27 | 69 | 45 | 23 | 65 | 40 | 21 | 52 | 30 | 25 | 68 | 42 | 25 | 70 | 43 |
| 2 | 23 | 61 | 34 | 27 | 68 | 43 | 34 | 81 | 60 | 30 | 76 | 52 | 29 | 63 | 44 | 38 | 79 | 58 | 36 | 78 | 57 |
| 3 | 33 | 76 | 53 | 33 | 77 | 54 | | | | | | | 40 | 76 | 60 | | | | | | |
| 4 | | | | | | | | | | | | | 51 | 84 | 68 | | | | | | |

approximation of F around its maximum. The point at which interpolation polynomial reaches a local maximum is considered the estimated age of that person. Fig. 4 presents an example of the age estimation process for a person with the following scoring:

PUSA = 1, PUSB = 1, PUSC = 1, SSPIA = 1, SSPIB = 2, SSPIC = 1 and SSPID = 1,

using intervals from Table 3.

Further improvement of this model can be achieved by taking into account the width of the individual intervals, preferring narrow intervals, In our model we used weighting that is inversely proportional to interval width, where the interval width covers about 80% of the population.

### 3.5. K nearest neighbors (KNN)

KNN is a simple algorithm that stores all available data and estimates the output value of new observations based on a similarity measure. In this case, the age is estimated as the average age of its K nearest neighbors, i.e. the most similar individuals in

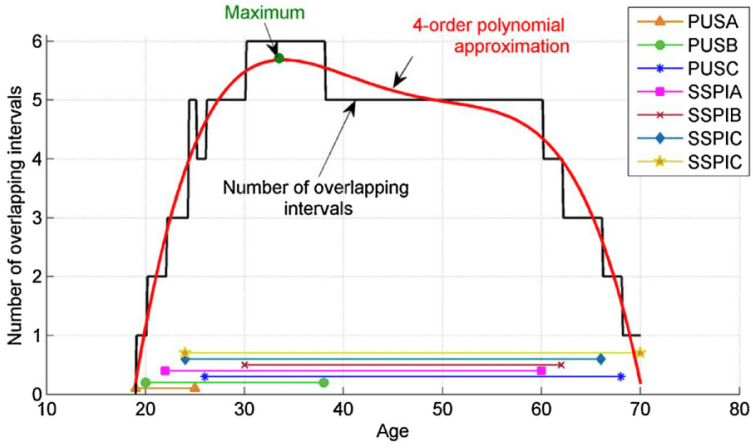


**Fig. 4.** Principle of the interval-based model.

the dataset. Similarity between individuals $p$ and $q$ is measured by a distance function in the form of Manhattan distance:

$$D = \sum_{i=1}^{N} |p_i - q_i|.$$

The estimated age is computed as the average of the $K$ nearest neighbors according to the following equation:

$$Age_{Estimated} = \Sigma_{i=1}^{K} p_i / K.$$

### 3.6. Artificial Neural Network (ANN)

We used multilayer perceptron (MLP), a feed-forward Artificial Neural Network that consists of multiple layers of nodes with each layer fully connected to the next one. Each node in the network after the input layer is a neuron with a non-linear activation function. We used single hidden layer network consisting of seven nodes in the input layer (corresponding to individual variables), four neurons in the hidden layer and one neuron in output layer. The total number of connections is 32. Backpropagation is used as the learning algorithm. It determines the value/weight of all connection between nodes so that the whole network can perform a specific task. Neurons in the hidden layer are all sigmoid; the output neuron is a linear unit. Fig. 5 shows an example of a neural network for age estimation.

### 3.7. Decision tree

The decision tree approach is a non-parametric supervised learning method which builds the model in the form of a tree structure. It breaks down a training dataset into smaller and smaller subsets. Splitting stops when all elements belong to a single class. The final result is a tree with decision nodes and leaf nodes providing various multi-linear functions. To reduce the complexity of the output model and to improve the robustness of the model, pruning was performed.

We performed both the regression and the classification experiments using CART decision trees on the whole dataset.

### 3.8. M5 tree

The M5 tree approach combines a conventional decision tree with the possibility of linear regression functions at the leaves.

### 3.9. Probabilistic models

Probabilistic age-at-death estimation has acquired special relevance because it is important to quantify and visualize the uncertainty associated with each estimate [46,63]. From a probabilistic point of view, age-at-death ($y$) estimated from skeletal indicators ($x$) can be defined as:

$$f(y|x) = \frac{f(x|y)f(y)}{f(x)} = \frac{f(x|y)f(y)}{\int f(x|y)f(y)dy}$$

In this study we fitted probabilistic models without modeling conditional dependencies [64,65], with the No Dependence Estimator (NDE) and with approximate conditional dependence modeling via the Averaged One-Dependence Estimator (AODE) algorithm [66].

The advantage of probabilistic models in age-at-death estimation is their ability to learn and predict with missing values, avoid the bias of projecting the training data age structure in target cases and compute both point-estimate and predictive intervals for each individual.

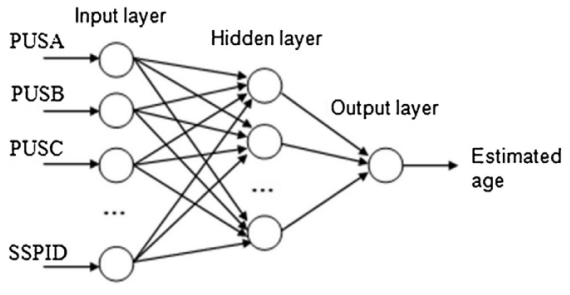


**Fig. 5.** Neural network for age estimation.

## 4. Results

The two-sample Kolmogorov–Smirnov test did not reject the null hypothesis (males and females are from the same continuous distribution) at the 5% significance level. The asymptotic p-value was 0.45.

Table 3 shows the minimum, maximum and median values of age for each variable and scoring after the lower and upper 10% of data were removed. For instance, the median age of the population having a score of 2 for the variable PUSA is 34 years, and 80% of that population is between 23 and 61 years old. The data are presented together for males and females, since no differences were observed in study of Buk et al. [51].

The accuracies of individual models were expressed by Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE), separately for females, males and the pooled sample. Although we report the results for both sexes, only those for the pooled sample are considered important according to the results of Buk et al. [51]. In the case of classification trees, the mean accuracy of the pooled sample was used to express the success rate of the classification.

Both RMSE and MAE measure the average magnitude of the errors in a set of predictions, but RMSE is the square root of the average of squared differences between a prediction and an actual observation whereas MAE is a linear scoring rule. MAE and RMSE express the average model prediction error, can range from zero to infinity and are negatively oriented (lower values are better) [67,68]. The Root Mean Squared Error is always greater than the Mean Absolute Error, which depends on sample size, corresponding with the results of Willmott and Matsuura [67]. MAE and RMSE are defined as:

$$MAE = \frac{1}{n}\sum_{j=1}^{n} |y_j - \hat{y}_j|$$

and

$$RMSE = \sqrt{\frac{1}{n}\sum_{j=1}^{n} \left(y_j - \hat{y}_j\right)^2}.$$

Table 4 contains results in years for both MAE and RMSE for all models, separately for females, males and the pooled sample. Results of MAE for single models were between 9.3 and 11.3 years in the case of males, between 10.3 and 12 years for females, and for the pooled sample values ranged between 9.7 and 11 years. Similarly, RMSE ranged between 11.6 and 14 years for males, between 12.7 and 14.7 years for females, and between 12.1 and 14.2 years for the pooled sample. The minimum values of MAE and RMSE for the pooled sample were reached by the Multi-linear regression model (9.7 and 12.1 years, respectively), followed by the Collapsed regression model (9.9 years and 12.2 years, respectively), whereas the maximum values were estimated by the Artificial

**Table 4**
Comparison of selected models which estimate the age as a number with use of 5-fold cross-validation.

| | Males | | Females | | Pooled | | Model is user friendly |
|---|---|---|---|---|---|---|---|
| | MAE (years) | RMSE | MAE (years) | RMSE | MAE (years) | RMSE | |
| Collapsed regression model | 9.3 | 11.6 | 10.4 | 12.7 | 9.9 | 12.2 | Yes |
| Multi-linear regression | 9.3 | 11.6 | 10.3 | 12.7 | 9.7 | 12.1 | Yes |
| Interval-based model | 10.3 | 12.9 | 12.0 | 14.7 | 10.8 | 13.3 | Yes |
| K Nearest Neighbors, K = 3 | 9.9 | 12.6 | 10.9 | 13.6 | 10.1 | 12.8 | No |
| Artificial Neural Network | 11.3 | 14.0 | 11.4 | 14.5 | 11.0 | 14.2 | No |
| Regression tree | 10.4 | 12.9 | 11.1 | 14.0 | 10.3 | 12.9 | Yes |
| M5 tree[a] | 9.3 | 11.6 | 10.3 | 12.7 | 9.7 | 12.1 | Yes |
| NDE[b] | 9.8 | 12.8 | 10.9 | 13.9 | 10.4 | 13.3 | No |
| AODE[b] | 9.8 | 12.7 | 10.6 | 13.6 | 10.2 | 13.2 | No |

MAE — Mean Absolute Error; RMSE — Root Mean Squared Error, numerical values in years.
[a] M5 tree collapsed into one leaf, thus it becomes equivalent to Multi-linear regression.
[b] Reference sample age distribution used as the prior distribution.

Neural Network (11 and 14.2 years, respectively). The M5 tree model turned out to be identical with the Multi-linear regression model after only one leaf was generated instead of a tree.

Furthermore, the usability of the models was evaluated.

Table 5compares the mean accuracy of five classification tree models for the pooled sample. The accuracy ranged between 30.7 and 72.3%, where the lowest accuracy was reached by Tree 3, a classification tree using both indicators classifying up to 10 years intervals, and the highest accuracy was reached by Tree 4, a classification tree using only the PUSx indicator, classifying into three age intervals. Trees number 1 (both indicators, age interval <30, 30–60, >60 years) and 2 (both indicators, age interval <30, 30–50, >50 years) reached a mean accuracy 69% and 68.6%, respectively. Classification tree number 5 (using only the SSPIx indicator, classifying into four age intervals) had a mean accuracy of 58.6%.

Below are the results for all the models listed in separate chapters. MAE and RMSE values for all models are shown in Table 4.

### 4.1. Collapsed regression model

We used the Trust-Region-Reflective Least Squares Algorithm implemented in the Matlab fitting tool to find the coefficients of the model. According to the Collapsed regression model, the age of the individual can be estimated as

$$Age_{Estimated} = a_1 * exp(a_2 * CUMM),$$

where $a_1 = 9.36$ and $a_2 = 0.1187$ for males, $a_1 = 10.79$ and $a_2 = 0.1107$ for females, and $a_1 = 10.16$ and $a_2 = 0.114$ for the pooled dataset, respectively. CUMM is defined as the sum of all individual scores (see Section 3.2). The 95% confidence bounds for coefficients $a_1$ and $a_2$ are: (7.807, 10.91) and (0.1077, 0.1297) for males; (9.046, 12.54) and (0.09978, 0.1216) for females; (8.982, 11.33) and (0.1063, 0.1217) for the pooled dataset.

### 4.2. Multi-linear regression

We used the data mining toolkit provided by the machine learning software Weka to find coefficients of Multi-linear regression model for males, females and the pooled dataset. The final models are the following:

Males:

$$Age_{Estimated} = 5.1 * PUSA + 4.1 * PUSB + 6.1 * PUSC + 8.3 * SSPIB + 4.2 * SSPIC + 5.9 * SSPID - 17.8$$

Females:

$$Age_{Estimated} = 4.2 * PUSA + 4.1 * PUSB + 6.2 * PUSC + 7.7 * SSPIB + 5.1 * SSPIC + 4.9 * SSPID - 12.9$$

Pooled dataset:

$$Age_{Estimated} = 4.5 * PUSA + 4.2 * PUSB + 6.1 * PUSC + 8.0 * SSPIB + 4.5 * SSPIC + 5.3 * SSPID - 14.8$$

### 4.3. Interval-based model

The interval-based model was implemented in Matlab. The integral part of this model is presented in Table 6, and the procedure is described in Section 3.4.

### 4.4. K nearest neighbors (KNN)

The KNN model includes the complete database of all 941 samples (the database is an essential part of the model) and is evaluated by the algorithm. Therefore, there is no simple formula that can be presented as a result, in contrast to the Collapsed regression model or the Multi-linear regression model.

### 4.5. Artificial Neural Network (ANN)

The neural network for the pooled dataset is described in Tables 7 and 8, where Table 7 describes connections from the input layer to four neurons in the hidden layer, and Table 8 describes connections from the hidden layer to one neuron in the output layer. The total number of connections is 32. These tables determine the weight of all connections between nodes and threshold values. Both the input attributes and the output of the network are normalized.

**Table 5**
Comparison of classification tree models which assign examined person into specific class, pooled sample, depth 4.

| | Tree 1 (Fig. 6) | Tree 2 (Fig. 7) | Tree 3 (Fig. 8) | Tree 4 (Fig. 9) | Tree 5 (Fig. 10) |
|---|---|---|---|---|---|
| Mean accuracy (%) | 69.0 | 68.6 | 30.7 | 72.3 | 58.6 |

Tree 1–both indicators, age interval <30, 30–60, >60 years; Tree 2 — both indicators, age interval <30, 30–50, >50 years; Tree 3 — both indicators, age interval 10-years intervals; Tree 4–PUSx indicator, age interval <30, 30–40, >40 years; Tree 5–SSPIx indicator, age interval <30, 30–40, 40–50, >50 years.

**Table 6**
Minimum and maximum values for 80% coverage of the data for each variable and scoring (in years).

| | Score | PUSA | | PUSB | | PUSC | | SSPIA | | SSPIB | | SSPIC | | SSPID | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | min | max | min | max | min | max | min | max | min | max | min | max | min | max |
| Males | 1 | 19 | 25 | 20 | 38 | 26 | 68 | 22 | 65 | 21 | 40 | 24 | 66 | 24 | 69 |
| | 2 | 23 | 57 | 27 | 68 | 35 | 80 | 32 | 76 | 30 | 62 | 38 | 79 | 36 | 78 |
| | 3 | 34 | 76 | 34 | 78 | | | | | 42 | 78 | | | | |
| | 4 | | | | | | | | | 51 | 81 | | | | |
| Females | 1 | 20 | 29 | 20 | 29 | 27 | 69 | 23 | 65 | 22 | 54 | 27 | 69 | 27 | 70 |
| | 2 | 23 | 63 | 24 | 66 | 32 | 81 | 29 | 74 | 28 | 66 | 38 | 81 | 33 | 77 |
| | 3 | 32 | 76 | 31 | 77 | | | | | 38 | 74 | | | | |
| | 4 | | | | | | | | | 49 | 87 | | | | |
| Pooled | 1 | 20 | 29 | 20 | 34 | 27 | 69 | 23 | 65 | 21 | 52 | 25 | 68 | 25 | 70 |
| | 2 | 23 | 61 | 27 | 68 | 34 | 81 | 30 | 76 | 29 | 63 | 38 | 79 | 36 | 78 |
| | 3 | 33 | 76 | 33 | 77 | | | | | 40 | 76 | | | | |
| | 4 | | | | | | | | | 51 | 84 | | | | |

### 4.6. Decision tree

Six decision trees were computed (5 classification trees and 1 regression tree) in Mathematica. Figs. 6–8 present decision trees where both skeletal indicator were used and the classification was performed within the following age intervals: <30, 30–60 and >60 years (Tree 1, Fig. 6), <30, 30–50, >50 years (Tree 2, Fig. 7) and 10-years intervals from 10–100 years (Tree 3, Fig. 8). Fig. 9 (Tree 4) presents a classification with the use of only one skeletal indicator, pubic symphysis, within three age intervals (<30, 30–40, >40 years) and Fig. 10 (Tree 5) was performed only for the auricular surface, within four age intervals (<30, 30–40, 40–50, >50 years). Fig. 11 depicts the only regression tree calculated with maximum depth limited to 4. In the case of classification trees, entropy is used to calculate the homogeneity of samples.

### 4.7. Probabilistic models

Probabilistic age estimation, especially by means of kernel density estimation methods (as performed here), is computationally expensive, and in multivariate cases it is hard or even impossible to represent the final model in a closed formula or a table, unlike other regression models presented in this paper.

Although this approach is not as straightforward as a closed-formula regression model, a simple web app was developed, available at https://dsnavega.shinyapps.io/koterova-etal/, in order to make the model easy to use for other researchers. Fig. 12 illustrates an age estimate using this approach.

## 5. Discussion

The main idea of the present study was to apply different computational approaches to visually scored data (two variables on the hip bone) to achieve more accurate age-at-death estimation. We had at our disposal a large dataset consisting of nine different samples. The same dataset, or at least part of it, was used in three previous studies. In the original paper of Schmitt et al. [28], a new scoring system for the auricular surface and the pubic symphysis of the hip bone was developed. Subsequently, the methodology proposed by Lucy et al. [69] was utilized to classify individuals into age categories. The results of Schmitt indicate that the combination of those two indicators does not perform better than the auricular surface alone.

**Table 7**
Parameters of synapsis and threshold values of sigmoid nodes in the hidden layer.

| | PUSA | PUSB | PUSC | SSPIA | SSPIB | SSPIC | SSPID | Threshold |
|---|---|---|---|---|---|---|---|---|
| Node 1 | 1.21 | 0.56 | 0.99 | −0.95 | 3.37 | 0.27 | 0.82 | −3.71 |
| Node 2 | −2.62 | −2.62 | 2.99 | −0.46 | 0.96 | −1.45 | −1.83 | −7.23 |
| Node 3 | 1.46 | −3.15 | 1.24 | −3.03 | 2.41 | 0.81 | 3.66 | −7.43 |
| Node 4 | −1.03 | −0.61 | −0.72 | 0.15 | −1.75 | −0.52 | −0.33 | −1.25 |

**Table 8**
Parameters of synapsis and threshold values of linear output node.

| | Node 1 | Node 2 | Node 3 | Node 4 | Threshold |
|---|---|---|---|---|---|
| Output node | 0.44 | −1.39 | −0.91 | −0.59 | −0.11 |

Ten years later, Martins et al. [65] used four skeletal series from this dataset and applied a modification of the empirical method (based on Bayesian decomposition with kernel smoothing) adjusted by Lucy et al. [64] to reduce the bias of the estimates. However, Martins et al. admitted that even though the statistical approach they used has certain advantages, its reliability remains similar to a previous study of Schmitt (2002). Interestingly, Martins et al. stressed that information about sex significantly improves the estimates. This contradicts the results of Buk et al. [51], who used the same original Schmitt's dataset supplemented by a few individuals. He concluded that knowledge of sex is not important in age classification, which is in agreement with the conclusions of other authors [7,34] This inconsistency may be given by the extended sample in the study of Buk and co-workers.

In their study several data mining methods – Feed-Forward Networks (FFN), Radial Basis Functions (RBF), Learning Vector Quantization (LVQ) and the Group of Adaptive Models Evolution (GAME) method – were used. They concluded that accurate and reliable age assessment is possible only to three broad age categories.

Since the dataset of Buk and co-workers is already complete, and thus the same as ours, we accepted the result that the sex variable is not significant for age estimation. Nevertheless, one should always test their own dataset since differences between sexes are possible [65].

Different mathematical models were applied to achieve the highest possible accuracy of age estimation, ideally to ten-year age intervals. Nevertheless, the results of RMSE and MAE analyses are self-explanatory, as is mean accuracy in the case of decision tree models. The values of MAE and RMSE were only slightly different from each other and ranged between 9.3 and 12 years and 11.6 and 14.7 years, respectively. Multi-linear regression followed by the Collapsed regression model performed the best. In the case of decision trees, the range of accuracy was much greater (30.7–72.3%). As could be expected, the lowest accuracy was reached by the tree classifying individuals into 10-years age intervals. Currently we are still unable to accurately and reliably classify

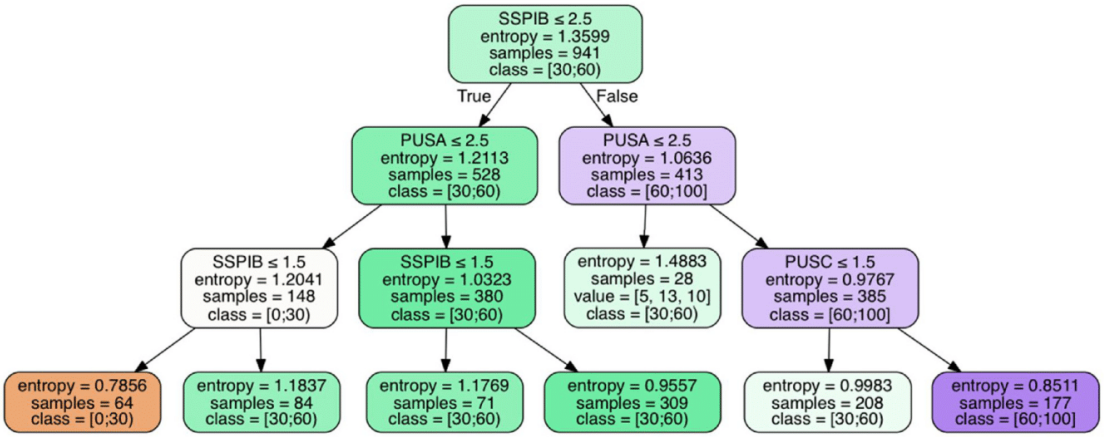


**Fig. 6.** Tree 1. Whole dataset, both indicators, age intervals under 30 years, 30–60 years, over 60 years.

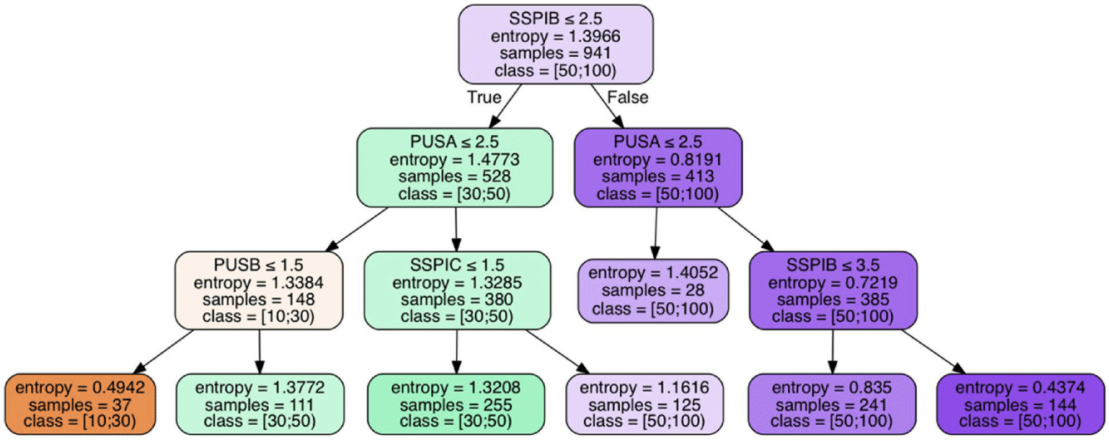


**Fig. 7.** Tree 2. Whole dataset, both indicators, age intervals under 30 years, 30–50 years, over 50 years.

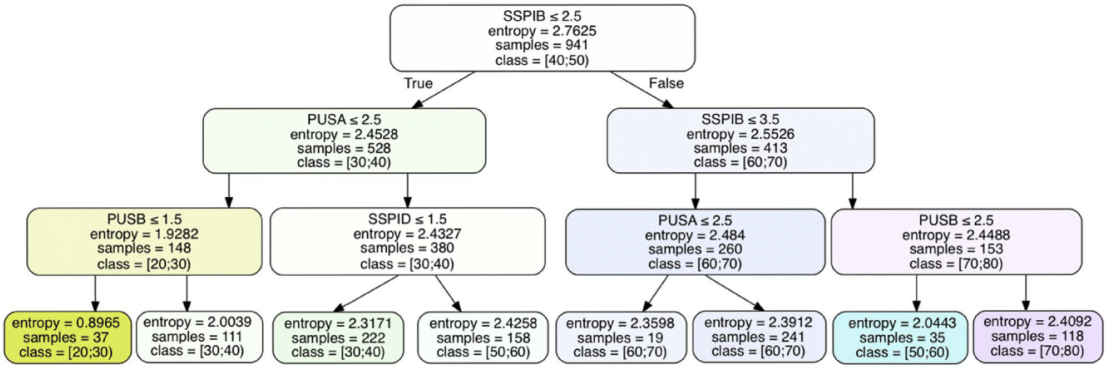


**Fig. 8.** Tree 3. Whole dataset, both indicators, age 10 years intervals (10–20, 20–30, 30–40, 40–50, 50–60, 60–70, 70–80, 80–90, 90–100).

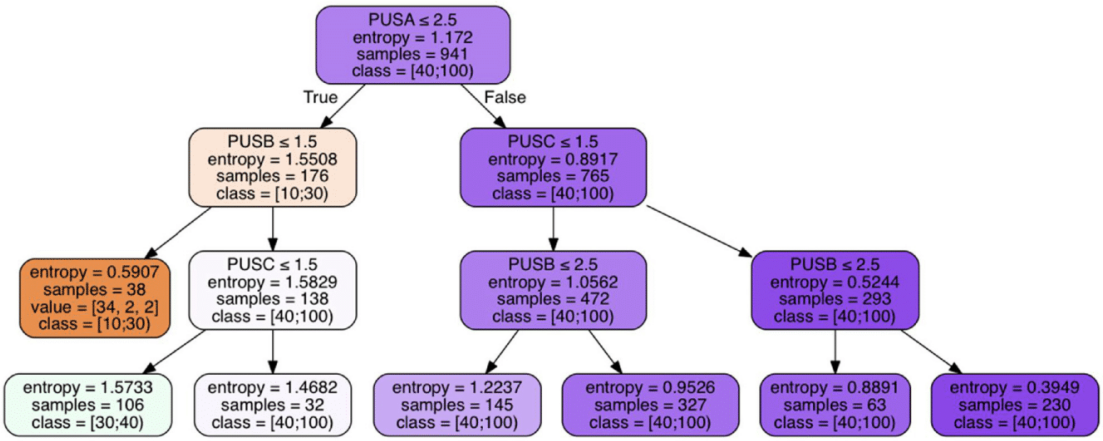


**Fig. 9.** Tree 4. Whole dataset, only PUSx indicator, age interval under 30 years, 30–40 years, over 40 years.

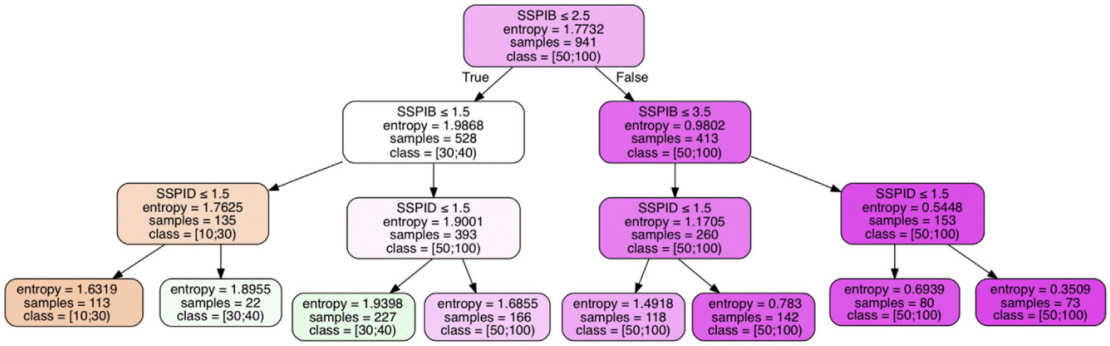


**Fig. 10.** Tree 5. Whole dataset, only SSPIx indicator, age interval under 30 years, 30–40 years, 40–50 years, over 50 years.

into 10-years intervals. Greater accuracy was reached by Tree 4, where only the PUSx indicator was involved. In this case, it is possible to classify with 72.3% accuracy into three age intervals. Two of them are closed (up to 30 years, 30–40 years) and the third is an open interval (>40 years). This corroborates previous findings that the pubic symphysis indicates age up to 40–45 years [16,56]. In the current situation, such three age intervals are the only compromise if 10-year intervals can be accepted, but only up to 40 years of age. It is technically impossible and biologically unjustifiable to expect a single skeletal indicator to capture age-related changes in the full range of adult ontogenesis. Data of the type used in our study contain a certain amount of noise, which is caused by multiple factors affecting scoring of senescence indicators. In terms of accuracy, probabilistic models behave like the other methods presented here. This was expected, as the prior for age-at-death was extracted from training data and that the same prior is implicit for the other methods presented here. The main advantage of this approach is that it allows to compute a case-specific posterior distribution of age-at-death based on a given observed skeletal morphology and, if necessary, specify another prior distribution over age-at-death.

Another very important question for anthropologists using such a model during fieldwork is how usable it is. Five models were evaluated as user friendly (Collapsed regression model, Multi-linear regression, Interval-based model, Regression tree and M5 tree), four were evaluated as more complicated (K nearest neighbors, Artificial Neural Network, NDE, AODE). However, in the case of probabilistic models (NDE and AODE), an easy-to-use web app has been developed. It should be stressed that the models with the lowest error rates are also considered user-friendly.

We have to admit that the objective of the present study was not met. We have obtained similar results as previous studies [51,65]. Non-parametric methods used in our study to analyze the same dataset do not perform better and we did not achieve more accurate classification.

We suppose that problems in estimating the age of adult individuals are caused by several factors:

1) subjectivity of the evaluation of morphological features; most of approaches are based on visual evaluation of surface changes (e.g. [16,46,70,71])
2) some factors which influence the senescence process (e.g. activity, body mass, body size, osteoarthritis, obesity) have been proven to have an effect (e.g. [11,72–75])
3) the fact that single indicators are used for the whole duration of adulthood (e.g. [7,27,32])
4) population specificity of the aging process (e.g. [51,76])
5) and, finally, variability of the aging process, which is simply too great [65] and unpredictable.

Based on our knowledge and experience from previous research, age-at-death estimation can be improved somewhat by taking the following steps:

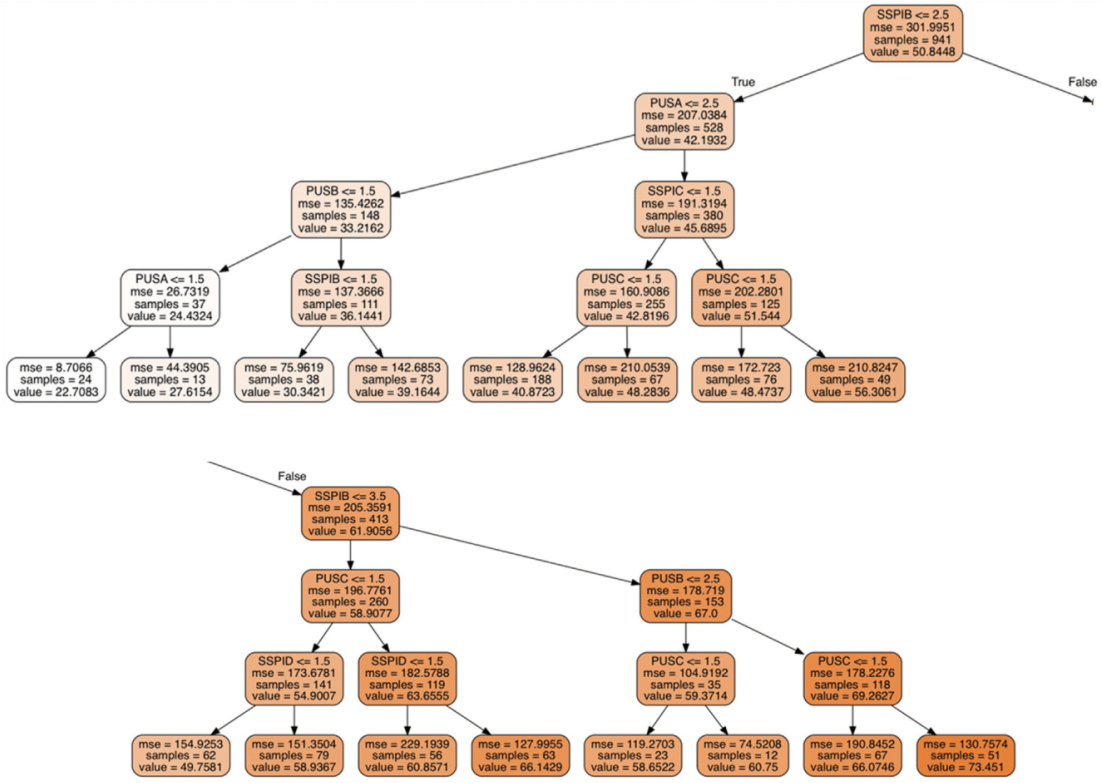


**Fig. 11.** Regression tree on the whole dataset with maximum depth limited to 4.

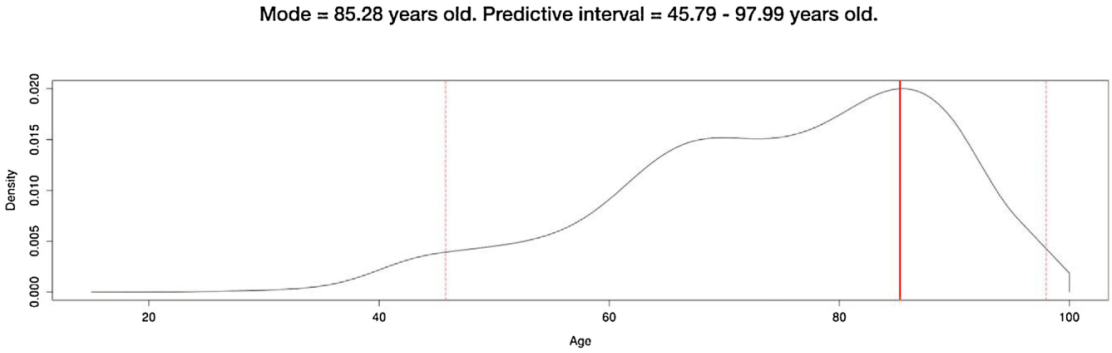


**Fig. 12.** Conditional density plot obtained by applying a bayesian age prediction method.

a) Substitute subjective evaluation of morphological changes by mathematical quantification of surface changes.

Future research should therefore focus on the mathematical quantification of surface changes. Recent studies on the pubic symphysis and on the auricular surface [49,70,77,78] seem very promising. In the case of the acetabulum, a method evaluating shape variability by means of geometric morphometrics was created in order to reduce the subjectivity of existing methods [36].

b) Avoid using only one age indicator for the whole lifespan.

Nowadays, the assumption that a single age indicator can reflect the whole age range and thus provide accurate estimations is considered incorrect. For example, degenerative changes on the pubic symphysis are observed only up to 40 years of age [16,51,56], after which the pubic symphysis does not show any changes usable for age estimation. Given that fact, several studies use the pubic symphysis only up to approximately 40 years of age (e.g. [56,79–81]). In our opinion, a very good approach has been proposed by Baccino and co-workers [82–84]. Their so-called Two-step procedure takes into account the chronological combination of age indicators with regard to in which age interval they predict best. The Two-step procedure (TSP, or the Two-step strategy) combines two methods — Suchey and Brooks [27], and Lamendin [85,86]. Each of them evaluates age best in different age categories: Suchey and Brooks up to 40 years of age (it correspond with the

first three stages of Suchey–Brooks), and Lamendin between 40 and 60 years of age. However, one huge disadvantage of TSP is that Lamendin's part is not applicable in bioarchaeology.

On the contrary, multi-indicator approaches, which simultaneously use several indicators, are expected to provide more accurate estimations [39–42,79,80]. However, these approaches, as they have been applied up until now, provide even broader age intervals [50] and lower accuracy. Yet, the main problem with multi-indicator methods is that they have not yet been properly standardized as to how to combine large number of skeletal indicators.

We believe that a chronological combination of age indicators has the right prerequisites to advance age estimation. This is also indicated by the results of Calce et al. [72] who confirmed differences in aging between different pelvic joints. The auricular surface exhibited the earliest age changes showed, the slowest changes occurred at the acetabulum, and the pubic symphysis showed the fastest age changes. These results are in concordance with some previous studies; for example, the acetabular area may be useful for the age estimation of older adult individuals [32,34,35].

c) Create multinational standards that will reduce population specificity.

The population specificity of the aging process causes the methods proposed for one population may not provide accurate results for another. At the same time, one has to be aware of globalization and the movement of people around the world, which is very topical lately [81,87]. One cannot be sure about the population affinity of unknown skeletal remains. The rate of misclassification in the assessment of ancestry is relatively high [88], although the research is still ongoing [53,89,90]. The only possible solution is to use a multi-population sample of known origin, as is the case in the present study.

Multi-population samples are considered [91] and proven to significantly improve different aspects of the biological profile. Therefore, according to several studies, some aspects exhibit greater accuracy and precision when robust methods based on several populations rather than a single one are used. That is because more variability is taken into account. Those pooled models performed better also for other aspects of the biological profile than sex determination [92,93] or the assessment of stature [94,95]. We suppose this is also applies to the estimation of age-at-death and some preliminary results support this idea [96].

Such are the prospects for further research towards greater accuracy and reliability of age-at-death estimation from adult human remains.

## 6. Conclusion

The presented manuscript compared nine different mathematical approaches in order to reach more accurate and precise age-at-death estimation. The best accuracy provide "Multilinear regression model" and "Collapsed regression model". The multilinear regression model gives MAE of 9.7 and RMSE of 12.1 for pooled dataset, respectively. The Collapsed regression model gives both MAE and RMSE about 0.1–0.2 higher. The best accuracy (30.7%) among decision trees models provide decision tree taking both indicators (PUSx, SSPIx) and 10-years intervals. We believe that the only way to improve the accuracy of age estimation is to avoid visual assessment of surface changes, which is highly subjective and does not capture the diversity of morphological expression. Thus, future research should focus on the objectification of methods (construction of 3D bone models, mathematical quantification of surface changes) and the classification of results with the use of machine learning methods.

## Acknowledgements

We would like to thank Dr. Aurore Schmitt from the Université de la Méditerranée – Secteur Nord, France for providing her scored dataset, to Grant Agency of the Czech Republic, research grant GAČR No. 17-01878S and to Mr. Fred Rooks for the linguistic corrections.

## References

[1] E. Cunha, E. Baccino, L. Martrille, F. Ramsthaler, J. Prieto, Y. Schuliar, N. Lynnerup, C. Cattaneo, The problem of aging human remains and living individuals: a review, Forensic Sci. Int. 193 (2009) 1–13.
[2] J.L. Boldsen, G.R. Milner, L.W. Konigsberg, J. Wood, Transition analysis: a new method for estimating age from skeletons, in: R. Hoppa, J. Vaupel (Eds.), Paleodemography: Age Distributions from Skeletal Samples, Cambridge University Press, Cambridge, 2002, pp. 73–106.
[3] M. Iscan, M. Steyn, The Human Skeleton in Forensic Medicine, Third ed., Charles C Thomas, Springfield, 2013.
[4] E. Baccino, E. Cunha, C. Cattaneo, Aging the dead and the living, in: J. Siegel, P. Saukko, M. Houck (Eds.), Encyclopedia of Forensic Sciences, Academic Press, Waltham, 2013, pp. 42–48.
[5] J. Buckberry, The (mis)use of adult age estimates in osteology, Ann. Hum. Biol. 42 (2015) 323–331.
[6] A. Cappella, M. Cummaudo, E. Arrigoni, F. Collini, C. Cattaneo, The issue of age estimation in a modern skeletal population: are even the more modern current aging methods satisfactory for the elderly? J. Forensic Sci. 62 (2017) 12–17.
[7] J.L. Buckberry, A.T. Chamberlain, Age estimation from the auricular surface of the ilium: a revised method, Am. J. Phys. Anthropol. 119 (2002) 231–239.
[8] S. Mays, The effect of factors other than age upon skeletal age indicators in the adult, Ann. Hum. Biol. 42 (2015) 332–341.
[9] G.E. Berg, Pubic bone age estimation in adult women, J. Forensic Sci. 53 (2008) 569–577.
[10] S.C. Zapico, D.H. Ubelaker, Applications of physiological bases of ageing to forensic sciences. Estimation of age-at-death, Ageing Res. Rev. 12 (2013) 605–617.
[11] C.E. Merritt, The influence of body size on adult skeletal age estimation methods, Am. J. Phys. Anthropol. 156 (2015) 35–57.
[12] C.E. Merritt, Inaccuracy and bias in adult skeletal age estimation: assessing the reliability of eight methods on individuals of varying body sizes, Forensic Sci. Int. 275 (2017), doi:http://dx.doi.org/10.1016/j.forsciint.2017.03.003 315.e1–315.e11.
[13] V.C. Campanacho, The Influence of Skeletal Size on Age-related Criteria from the Pelvic Joints in Portuguese and North American Samples, University of Sheffield, 2016.
[14] R.D. Hoppa, J.W. Vaupel, The Rostock manifesto for paleodemography: the way from stage to age, in: R.D. Hoppa, J.W. Vaupel (Eds.), Paleodemography: Age Distribution from Skeletal Sample, Cambridge University Press, Cambridge, 2002, pp. 1–8.
[15] L.W. Konigsberg, Multivariate cumulative probit for age estimation using ordinal categorical data, Ann. Hum. Biol. 42 (2015) 368–378.
[16] N. Márquez-Grant, An overview of age estimation in forensic anthropology: perspectives and practical considerations, Ann. Hum. Biol. 42 (2015) 308–322.
[17] C. Merritt, Testing the accuracy of adult skeletal age estimation methods: original methods versus revised and newer methods, Explor. Anthropol. 12 (2014) 102–119.
[18] M.Y. Iscan, S.R. Loth, R.K. Wright, Age estimation from the rib by phase analysis: white males, J. Forensic Sci. 29 (1984) 1094–1104.
[19] Y.M. Iscan, S.R. Loth, R.K. Wright, Metamorphosis at the sternal rib end: a new method to estimate age at death in white males, Am. J. Phys. Anthropol. 65 (1984) 147–156.
[20] M.Y. Iscan, S.R. Loth, R.K. Wright, Age estimation from the rib by phase analysis: white females, J. Forensic Sci. 30 (1985) 853–863.
[21] K.M. Hartnett, Analysis of age-at-death estimation using data from a new, modern autopsy sample — part II: sternal end of the fourth rib, J. Forensic Sci. 55 (2010) 1152–1156.
[22] C.G. Falys, D. Prangle, Estimating age of mature adults from the degeneration of the sternal end of the clavicle, Am. J. Phys. Anthropol. 156 (2015) 203–214.
[23] N.V. Passalacqua, Forensic age-at-death estimation from the human sacrum, J. Forensic Sci. 54 (2009) 255–262.
[24] R.S. Meindl, C.O. Lovejoy, Ectocranial suture closure: a revised method for the determination of skeletal age at death based on the lateral-anterior sutures, Am. J. Phys. Anthropol. 68 (1985) 57–66.
[25] W.T. Todd, Age changes in the pubic bone. I. The male white pubis, Am. J. Phys. Anthropol. 3 (1920) 285–334.
[26] T.W. Mckern, T.D. Stewart, Skeletal age changes in young american males, Technical Report EP 45, Quartermaster Research and Development Command, Natick, 1957.
[27] S. Brooks, J.M. Suchey, Skeletal age determination based on the os pubis: a comparison of the Acsádi–Nemeskéri and Suchey–Brooks methods, Hum. Evol. 5 (1990) 227–238.

[28] A. Schmitt, P. Murail, E. Cunha, D. Rougé, Variability of the pattern of aging on the human skeleton: evidence from bone indicators and implication on age at death estimation, J. Forensic Sci. 47 (2002) 1203–1209.
[29] H.M. Garvin, N.V. Passalacqua, Current practices by forensic anthropologists in adult skeletal age estimation, J. Forensic Sci. 57 (2012) 427–433.
[30] C.O. Lovejoy, R.S. Meindl, T.R. Pryzbeck, R.P. Mensforth, Chronological metamorphosis of the auricular surface of the ilium: a new method for the determination of adult skeletal age at death, Am. J. Phys. Anthropol. 68 (1985) 15–28.
[31] Y. Igarashi, K. Uesu, T. Wakebe, E. Kanazawa, New method for estimation of adult skeletal age at death from the morphology of the auricular surface of the ilium, Am. J. Phys. Anthropol. 128 (2005) 324–339.
[32] C. Rissech, G.F. Estabrook, E. Cunha, A. Malgosa, Using the acetabulum to estimate age at death of adult males, J. Forensic Sci. 51 (2006) 213–229.
[33] C. Rougé-Maillart, N. Jousset, B. Vielle, A. Gaudin, N. Telmon, Contribution of the study of acetabulum for the estimation of adult subjects, Forensic Sci. Int. (2007) 103–110.
[34] S.E. Calce, A new method to estimate adult age-at-death using the acetabulum, Am. J. Phys. Anthropol. 148 (2012) 11–23.
[35] M. San-Millán, C. Rissech, D. Turbón, New approach to age estimation of male and female adult skeletons based on the morphological characteristics of the acetabulum, Int. J. Leg. Med. (2016) 1–25.
[36] M. San-Millán, C. Rissech, D. Turbón, Shape variability of the adult human acetabulum and acetabular fossa related to sex and age by geometric morphometrics. Implications for adult age estimation, Forensic Sci. Int. 272 (2017) 50–63.
[37] T. Colard, B. Bertrand, S. Naji, Y. Delannoy, A. Bécart, Toward the adoption of cementochronology in forensic context, Int. J. Leg. Med. (2015) 1–8, doi:http://dx.doi.org/10.1007/s00414-015-1172-8.
[38] L.A. Meckel, The utility of dental cementum increment analysis for estimating season-of-death in naturally decomposed skeletons, Annual Meetings — American Association of Physical Anthropologists, New Orleans, 2017.
[39] D. Ferembach, I. Schwidetzky, M. Stloukal, Recommendations for age and sex diagnoses of skeletons, J. Hum. Evol. 9 (1980) 517–549.
[40] M.E. Bedford, K.F. Russell, C.O. Lovejoy, R.S. Meindl, S.W. Simpson, P.L. Stuart-Macadam, Multifactorial aging method using skeletons with known ages-at-death from the grant collection, Am. J. Phys. Anthropol. 91 (1993) 287–297.
[41] C.O. Lovejoy, R.S. Meindl, R.P. Mensforth, T.J. Barton, Multifactorial determination of skeletal age at death: a method and blind tests of its accuracy, Am. J. Phys. Anthropol. 68 (1985) 1–14.
[42] M. Houck, D. Ubelaker, D. Owsley, E. Craig, W. Grant, R. Fram, T. Woltanski, K. Sandness, The role of forensic anthropology in the recovery and analysis of Branch Davidian compound victims: assessing the accuracy of age estimations, J. Forensic Sci. 41 (1996) 796–801.
[43] M.E. Bedford, K.F. Russell, C.O. Lovejoy, R.S. Meindl, S.W. Simpson, P.L. Stuart-Macadam, Test of the multifactorial aging method using skeletons with known ages-at-death from the grant collection, Am. J. Phys. Anthropol. 91 (1993) 287–297.
[44] X. Chen, Z. Zhang, L. Tao, Determination of male age at death in Chinese Han population: using quantitative variables statistical analysis from pubic bones, Forensic Sci. Int. 175 (2008) 36–43.
[45] N.R. Shirley, P.A. Ramirez Montes, Age estimation in forensic anthropology: quantification of observer error in phase versus component-based methods, J. Forensic Sci. 60 (2015) 107–111.
[46] C. Rougé-Maillart, B. Vielle, N. Jousset, D. Chappard, N. Telmon, E. Cunha, Development of a method to estimate skeletal age at death in adults using the acetabulum and the auricular surface on a Portuguese population, Forensic Sci. Int. 188 (2009) 91–95.
[47] H. Biwasaka, K. Sato, Y. Aoki, H. Kato, Y. Maeno, T. Tanijiri, S. Fujita, K. Dewa, Three dimensional surface analyses of pubic symphyseal faces of contemporary Japanese reconstructed with 3D digitized scanner, Leg. Med. 15 (2013) 264–268.
[48] M. Obert, M. Seyfried, F. Schumacher, G.A. Krombach, M.A. Verhoff, Aging adult skull vaults by applying the concept of fractal geometry to high-resolution computed tomography images, Forensic Sci. Int. 242 (2014) 24–31, doi:http://dx.doi.org/10.1016/j.forsciint.2014.06.018.
[49] D. Stoyanova, B.F.B. Algee-Hewitt, D.E. Slice, An enhanced computational method for age-at-death estimation based on the pubic symphysis using 3D laser scans and thin plate splines, Am. J. Phys. Anthropol. 158 (2015) 431–440.
[50] G.R. Milner, J.L. Boldsen, Transition analysis: a validation study with known-age modern American skeletons, Am. J. Phys. Anthropol. 148 (2012) 98–110.
[51] Z. Buk, P. Kordik, J. Bruzek, A. Schmitt, M. Snorek, The age at death assessment in a multi-ethnic sample of pelvic bones using nature-inspired data mining methods, Forensic Sci. Int. 220 (2012) 294.e1–e9.
[52] J.T. Hefner, M.K. Spradley, B. Anderson, Ancestry assessment using random forest modeling, J. Forensic Sci. 59 (2014) 583–589.
[53] D. Navega, C. Coelho, R. Vicente, M.T. Ferreira, S. Wasterlain, E. Cunha, AncesTrees: ancestry estimation with randomized decision trees, Int. J. Leg. Med. 129 (2015) 1145–1153.
[54] D.K.B. Hufnagl, Age Estimation with Decision Trees: Testing the Relevance of 94 Aging Indicators on the William M. Bass Donated Collection, University of Tennessee, 2015.
[55] N.R. Langley, B. Dudzik, A. Cloutier, A decision tree for nonmetric sex assessment from the skull, J. Forensic Sci. (2017) 1–7, doi:http://dx.doi.org/10.1111/1556-4029.13534.
[56] B. Dudzik, N.R. Langley, Estimating age from the pubic symphysis: a new component-based system, Forensic Sci. Int. 257 (2015) 98–105.
[57] E. Cunha, S. Wasterlain, The Coimbra identified osteological collections, in: G. Grupe, J. Peters (Eds.), Skeletal Series and Their Socioeconomic Context, Verlag Leidorf, Munchen, 2007, pp. 23–33.
[58] G. Perreard Lopreno, Adaptation structurelle des os du membre supérieur et de la clavicule à l'activité, University of Geneva, 2007.
[59] A. Schmitt, Variabilité de la sénescence du squelette humain. Réflexions sur les indicateurs de l'âge au décès: à la recherche d'un outil performant, University of Bordeaux, 2001.
[60] R.P. Mensforth, B.M. Latimer, Hamann–Todd collection aging studies: osteoporosis fracture syndrome, Am. J. Phys. Anthropol. 80 (1989) 461–479.
[61] M.R. Dayal, A.D.T. Kegley, G. Štrkalj, M.A. Bidmos, K.L. Kuykendall, The history and composition of the Raymond A. Dart collection of human skeletons at the University of the Witwatersrand, Johannesburg, South Africa, Am. J. Phys. Anthropol. 140 (2009) 324–335.
[62] A. Schmitt, Une nouvelle méthode pour estimer l'âge au décès des adultes à partir de la surface sacro-pelvienne iliaque, Bull. Mem. Soc. Anthropol. Paris 17 (2005) 89–101.
[63] L.W. Konigsberg, N.P. Herrmann, D.J. Wescott, E.H. Kimmerle, Estimation and evidence in forensic anthropology: age-at-death, J. Forensic Sci. 53 (2008) 541–557.
[64] D. Lucy, R.G. Aykroyd, A.M. Pollard, Nonparametric calibration for age estimation author, Appl. Stat. 51 (2002) 183–196.
[65] R. Martins, P.E. Oliveira, A. Schmitt, Estimation of age at death from the pubic symphysis and the auricular surface of the ilium using a smoothing procedure, Forensic Sci. Int. 219 (2012) 287.e1–e7.
[66] G.I. Webb, J.R. Boughton, Z. Wang, Not so naive Bayes: aggregating one-dependence estimators, Mach. Learn. 58 (2005) 5–24.
[67] C.J. Willmott, K. Matsuura, Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance, Clim. Res. 30 (2005) 79–82.
[68] G. Li, J. Shi, On comparing three artificial neural networks for wind speed forecasting, Appl. Energy 87 (2010) 2313–2320.
[69] D. Lucy, R.G. Aykroyd, A.M. Pollard, T. Solheim, A bayesian approach to adult human age estimation from dental observations by Johanson's age changes, J. Forensic Sci . 41 (1996) 189–194.
[70] D.E. Slice, B.F.B. Algee-Hewitt, Modeling bone surface morphology: a fully quantitative method for age-at-death estimation using the pubic symphysis, J. Forensic Sci. 60 (2015) 835–843.
[71] C. Villa, J. Buckberry, C. Cattaneo, N. Lynnerup, Technical note: reliability of Suchey-Brooks and Buckberry-Chamberlain methods on 3D visualizations from CT and laser scans, Am. J. Phys. Anthropol. 151 (2013) 158–163.
[72] S.E. Calce, H.K. Kurki, D.A. Weston, L. Gould, The effects of osteoarthritis on age-at-death estimates from the human pelvis, Annual Meetings — American Association of Physical Anthropologists, New Orleans, 2017.
[73] J. Truesdell, The effect of lifestyle factors such as smoking, activity, menopause, and pregnancy on age estimation from the pubic symphysis: a study of 1,238 living volunteers, Annual Meetings — American Association of Physical Anthropologists, New Orleans, 2017.
[74] C.E. Merritt, Body size as a factor in skeletal age estimation: when size matters and how to deal with it, Annual Meetings — American Association of Physical Anthropologists, New Orleans, 2017.
[75] D.J. Wescott, J.L. Drew, Effect of obesity on the reliability of age-at-death indicators of the pelvis, Am. J. Phys. Anthropol. 156 (2015) 595–605.
[76] A. Schmitt, Age-at-death assessment using the os pubis and the auricular surface of the ilium: a test on an identified Asian sample, Int. J. Osteoarchaeol. 14 (2004) 1–6.
[77] D.K. Stoyanova, B.F.B. Algee-Hewitt, J. Kim, D.E. Slice, A computational framework for age-at-death estimation from the skeleton: surface and outline analysis of 3D laser scans of the adult pubic symphysis, J. Forensic Sci. 62 (2017) 1434–1444.
[78] C. Villa, J. Buckberry, C. Cattaneo, B. Frohlich, N. Lynnerup, Quantitative analysis of the morphological changes of the pubic symphyseal face and the auricular surface and implications for age at death estimation, J. Forensic Sci. 60 (2015) 556–565.
[79] E. Baccino, J.C. Tavernier, H. Lamendin, D. Frammery, R. Nossintchouk, J.F. Humbert, Recherche d'une méthode multifactorielle simple pour la détermination de l'age des cadavres adultes, J. Med. Leg. Droit. Med. 34 (1991) 27–33.
[80] E. Baccino, A. Zerilli, The two step strategy (TSS) or the right way to combine a dental (Lamendin) and an anthropological (Suchey–Brooks system) method for age determination, Am. Acad. Forensic Sci. N Y (1997) p. 150.
[81] E. Baccino, L. Sinfield, S. Colomb, T.P. Baum, L. Martrille, Technical note: the two step procedure (TSP) for the determination of age at death of adult human remains in forensic cases, Forensic Sci. Int. 244 (2014) 247–251.
[82] H. Lamendin, Appréciation d'age par la méthode de Gustafson <<simplifiée>>, Chir. Dent. Fr. 427 (1988) 205–214.
[83] H. Lamendin, E. Baccino, J.F. Humbert, J.C. Tavernier, R.M. Nossintchouk, A. Zerilli, A simple technique for age estimation in adult corpses: the two criteria dental method, J. Forensic Sci. 37 (1992) 1373–1379 JFSCA.
[84] G. Acsadi, I. Nemeskeri, History of Human Life Span and Mortality, Akademiai Kiado, Budapest, 1970.
[85] R. Prodhan, D. Ubelaker, Evaluation of three methods of age estimation from human skeletal remains (Suchey–Brooks, Lamendin, and two-step strategy), Proceedings of the American Academy of Forensic Sciences, Seattle, 2006.

[86] E. L'Abbé, M. Steyn, The establishment and advancement of forensic anthropology in South Africa, in: D.C. Dirkmaat (Ed.), A Companion to Forensic Anthropology, Wiley-Blackwell, Malden, 2012, pp. 626–638.
[87] J. Bruzek, P. Murail, Methodology and reliability of sex determination from the skeleton, in: A. Schmitt, E. Cunha, J. Pinheiro (Eds.), Forensic Anthropology and Medicine. Complementary Sciences From Recovery to Cause of Death, Humana Press Inc., Totowa, 2006, pp. 225–242.
[88] B. Dudzik, R.L. Jantz, Misclassifications of hispanics using Fordisc 3.1: comparing cranial morphology in Asian and Hispanic populations, J. Forensic Sci. 61 (2016) 1311–1318.
[89] G. Scott, D. Navega, J. Coelho, E. Cunha, J. Irish, rASUDAS: a new method for estimating ancestry from tooth crown and root morphology, Annual Meetings — American Association of Physical Anthropologists, Atlanta, 2016.
[90] M. Katherine Spradley, R.L. Jantz, Ancestry estimation in forensic anthropology: geometric morphometric versus standard and nonstandard interlandmark distances, J. Forensic Sci. 61 (2016) 892–897.
[91] A. Koterova, J. Veleminska, J. Dupej, H. Brzobohata, A. Pilny, J. Bruzek, Disregarding population specificity: its influence on the sex assessment methods from the tibia, Int. J. Legal Med. 131 (2017) 251–261.
[92] P. Murail, J. Bruzek, F. Houët, E. Cunha, DSP: a tool for probabilistic sex diagnosis using worldwide variability in hip-bone measurement, Bull. Mem. Soc. Anthropol. Paris 17 (2005) 167–176.
[93] J. Brůžek, F. Santos, B. Dutailly, P. Murail, E. Cunha, Validation and reliability of the sex estimation of the human os coxae using freely available DSP2 software for bioarchaeology and forensic anthropology, Am. J. Phys. Anthropol. (2017) 1–10, doi:http://dx.doi.org/10.1002/ajpa.23282.
[94] C.B. Ruff, B.M. Holt, M. Niskanen, V. Sládek, M. Berner, E. Garofalo, H.M. Garvin, M. Hora, H. Maijanen, S. Niinimäki, K. Salo, E. Schuplerová, D. Tompkins, Stature and body mass estimation from skeletal remains in the European Holocene, Am. J. Phys. Anthropol. 148 (2012) 601–617.
[95] A. Hatza, S.D. Ousley, L.L. Cabo, Estimating stature when ancestry is unknown: what statistical methods work best? Proceedings — American Academy of Forensic Sciences, Orlando, 2015, pp. 152–153.
[96] J. Kim, Understanding population-specific age estimation using documented Asian skeletal samples, Annual meetings — American Association of Physical Anthropologists, New Orleans, 2017.

# Appendix A.3

Journal of Forensic and Legal Medicine 68 (2019) 101866

Contents lists available at ScienceDirect

Journal of Forensic and Legal Medicine

journal homepage: www.elsevier.com/locate/yjflm

Research Paper

## Impact of 3D surface scanning protocols on the *Os coxae* digital data: Implications for sex and age-at-death assessment

Anežka Kotěrová[a,*], Vlastimil Králík[b], Rebeka Rmoutilová[a,c], Lukáš Friedl[d,e], Pavel Růžička[b], Jana Velemínská[a], François Marchal[f], Jaroslav Brůžek[a,c]

[a] Department of Anthropology and Human Genetics, Faculty of Science, Charles University, Viničná 7, Prague, 128 43, Czech Republic
[b] Department of Mechanics, Biomechanics and Mechatronics, Faculty of Mechanical Engineering, CTU in Prague, Technická 4, Prague, 166 07, Czech Republic
[c] Laboratoire PACEA, UMR 5199, CNRS, Université Bordeaux, CS 50023, Pessac, 33615, France
[d] Department of Anthropology, University of West Bohemia, Plzeň, 30614, Czech Republic
[e] Interdisciplinary Center for Archaeology and Evolution of Human Behaviour (ICArEHB), Faculdade das Ciências Humanas e Sociais, Universidade do Algarve, Campus Gambelas, 8005-139, Faro, Portugal
[f] UMR 7268 ADES, Aix-Marseille University, EFS, CNRS, Faculté de Médecine Secteur Nord, 13344, Marseille Cedex 15, France

ARTICLE INFO

*Keywords:*
Laser scanning
Structured light technology
RedLux profiler
Biological profile
Age and sex estimation
*Os coxae*

ABSTRACT

The 3D imaging technologies have become of paramount importance for example in disciplines such as forensic anthropology and bioarchaeology, where they are being used more and more frequently. There are several new possibilities that they offer; for instance, the easier and faster sharing of data among institutions, the possibility of permanent documentation, or new opportunities of data analysis. An important requirement, however, is whether the data obtained from different scanning devices are comparable and whether the possible varying outputs could affect further analyses, such as the estimation of the biological profile. Therefore, we aimed to investigate two important questions: (1) whether 3D models acquired by two different scanning technologies (structured light and laser) are comparable and (2) whether the scanning equipment has an effect on the anthropological analyses, such as age-at-death estimation and sex assessment.

3D models of *ossa coxa* (n = 29) were acquired by laser (NextEngine) and structured light (HP 3D Structured Light Scanner PRO 2) scanners. The resulting 3D models from both scanners were subjected to age-at-death analyses (via the quantitative method of Stoyanova et al., 2017) and sex analyses (via Diagnose Sexuelle Probabiliste 2 of Brůžek et al., 2017). Furthermore, high quality scans of a small sample (n = 5) of pubic symphyseal surfaces with the RedLux Profiler device were acquired as reference surfaces to which the outputs from both scanners were compared. Small deviations between surfaces were more evident in more rugged surfaces (in areas of depression and protrusion). Even though small differences from the reference surfaces were found, they did not have a significant effect on the age and sex estimates. It never resulted in the opposite sex assignment, and no significant differences were observed between age estimates (with the exception of those with the TPS/BE model).

### 1. Introduction

The 3D imaging technologies, i.e. surface scanning (both, laser and structured light), as well as computed tomography (CT) scanners and micro-CT scanners have become widely used in anatomical research over the last few years.[1] All of these devices make it possible to obtain 3D models of desired objects. Surface scanning, contrary to CT scanning, offers several practical advantages, for instance, the higher portability of most devices, texture capture options, low cost, and rapid post-processing. Furthermore, they can be operated without certification since there is no radiation involved during the scanning process.[2–4] On the other hand, internal structures remain hidden.[4] Scanning technologies based on visible light have spread to many different disciplines: e.g. to forensic science,[5,6] anthropology and paleoanthropology,[2–4,7] anatomy and morphology,[8–11] and paleontology and archaeology.[8,12,13] They have allowed for new applications and brought additional advantages. It is now possible, for instance, to create and archive digital copies of skeletal remains as 3D models in the virtual environment. Such digital osteological collections are invaluable for researchers for several reasons. First, they represent the possibility

* Corresponding author. Department of Anthropology and Human Genetics, Charles University, Viničná 7, Prague 2, 128 43, Czech Republic.
*E-mail address:* koterova@natur.cuni.cz (A. Kotěrová).

https://doi.org/10.1016/j.jflm.2019.101866
Received 1 June 2019; Received in revised form 30 August 2019; Accepted 3 September 2019
Available online 05 September 2019
1752-928X/ © 2019 Elsevier Ltd and Faculty of Forensic and Legal Medicine. All rights reserved.

of permanent storage or documentation and conservation of bones, which are then accessible even though the real remains are no longer available or do not exist anymore. Virtual storage can serve not only as a repository, but also enables easier and faster sharing of data among researchers and institutions (e.g. Refs. 1,3,4,8,12,14–16). The reduced need for physical manipulation with human remains is certainly a great advantage[15] since the preservation of dry bones varies as they are constantly being used for scientific purposes.[17] Virtual models could be used for teaching and research, as well as for 3D printing, which can be used for exhibition purposes[10,18,19] or as demonstrative evidence in court.[20,21] Last but not least, digital technologies enable us to virtually reconstruct damaged skeletal remains, e.g. incompletely preserved fossil remains.[22–24]

Traditional methods of biological profile estimation that use a morphometric or visual approach are usually used for defining the biological profile of an individual in bioarchaeology as well as in forensic anthropology (e.g. Refs. 25,26). However, virtual bone models have become of paramount importance and are commonly used, for instance, to measure metric variables in order to estimate sex.[27,28] The consistency between the dimensions taken directly from dry bones and from their virtual representations,[28–31] and even from their printed replicas,[20] has been proven repeatedly. Sex and stature estimates via linear measurements taken from virtual data obtained with two different laser scanners showed only a very small deviation from traditionally obtained data in one recent study.[32] Moreover, bones in the virtual environment can undergo analyses that cannot be performed on dry bones, e.g. quantitative analyses of the surface with the use of geometric morphometrics.

For example, geometric morphometric tools are used for sex estimation[33–35] and ancestry assessment.[36,37] To estimate age-at-death, pelvic articulations (the pubic symphysis, the auricular surface, and the acetabulum) are often used since their surfaces undergo changes with aging and, recently, these surfaces have begun to be evaluated quantitatively.[38–42] Such analyses tend to be more objective than traditional visual methods based on scoring as they allow for the evaluation of the morphological variation of the skeletal remains independent of the human eye and experience of the researcher.[18]

Villa et al.[43] raised an important question as to whether the results of surface quantification used for age estimation from different laser scanners are comparable. The precision and repeatability of measurements among different scanning devices are necessary in order to ensure the reliability of the biological profile estimation methods. We have extended this question to another parameter of biological profile – sex estimation (besides age estimation) and, in contrast to the original study where only laser scanners were used, both laser and structured light scanners were compared.

The aims of the present study are twofold. The first is to compare 3D models of *os coxae* made by two different surface scanners (NextEngine laser scanner and HP 3D Structured Light Scanner Pro S2) with a reference sample derived from a RedLux Profiler device. Second, we aimed at assessing whether the scanning equipment has an effect on the age-at-death estimation (via the quantitative method of Stoyanova et al., 2017) and sex assessment (via DSP2 - Diagnose Sexuelle Probabiliste of Brůžek et al., 2017).

## 2. Material

The skeletal sample used in the present study originates from the cemetery of the second church in the Mikulčice settlement (9th–10th century AD), South Moravia, Czech Republic. This sample represents a medieval population of Central Europe that belonged to the Great Moravian Empire.[44] We used 18 adult individuals, of which 11 had well preserved *ossa coxa* on both sides, and seven individuals who had only the left or right bone well preserved. Altogether, 29 *ossa coxa* were used in this study. Only individuals with well-preserved articular surfaces (pubic symphysis and auricular surface) were selected. Even if we do not know the real age or sex of individuals, we can compare the resulting estimates with each other when they are derived from surface models digitized with different scanning technologies (HP 3D SLS, NextEngine and RedLux).

## 3. Methods

### 3.1. Digitization of skeletal material

A small sample of pubic symphyses (n = 5) was digitized with the use of a RedLux Profiler contactless metrology device (RedLux Ltd., Southampton, UK).[45,46] Samples were selected to include both smooth and significantly billowed symphyses. The scanned surface of key areas was utilized as a reference surface to compare the resolution and quality of scans obtained by commonly used scanners. The whole surface of *os coxae* (n = 29) was then digitized with two different scanning devices: the NextEngine 3D scanner Ultra HD and the HP 3D Structured Light Scanner PRO S2. The scanning process as well as the post-processing procedures of all three scanners are described in the following sections.

#### 3.1.1. RedLux

The pubic symphyses of five selected *ossa coxa* were digitized using the RedLux profiler, which is a device designed for very precise surface measurements using a confocal sensor. Due to limited space capacity resulting from the technical arrangement of the measuring device, the physical casts of these five pubic symphyses needed to be obtained. As a casting material, the two-component Addition Cure Molding Rubber was used, which is known under the trade designation MM242R.[47] This material shows negligible volumetric change (linear shrinkage of 0.09%) and excellent quality and accuracy of surface reconstruction. It is also necessary to ensure thorough venting, using a desiccator and vacuum pump so that no air bubbles affect the volumetric change. Venting was carried out for a minimum of 10 min at 150 mbar vacuum. The resulting symphyseal cast for one selected sample is shown in Fig. 1.

The casts were scanned using a highly accurate RedLux profiler device. The instrument was equipped with two high-precision movable linear axes and two rotary ones. The rotary sections carry the sample and the linear sections carry the sensor. With this sensor, the lens error, commonly known as chromatic aberration, is used to measure the distance to an object. By combining the sensor signal with the knowledge of the exact position of all 4 stages, 3D representation of the surface can be created. The instrument has the capability to measure the entire surface in a single procedure. The accuracy of the resulting point cloud is given by the resolution of 2 linear axes, the resolution of the two rotational axes and the resolution of the probe. The resolution of each linear axis is 100 nm, the resolution of each rotary axis is 10 arc seconds and the resolution of the probe is 20 nm. A detailed description of the

**Fig. 1.** Symphyseal cast of sample no. 390.

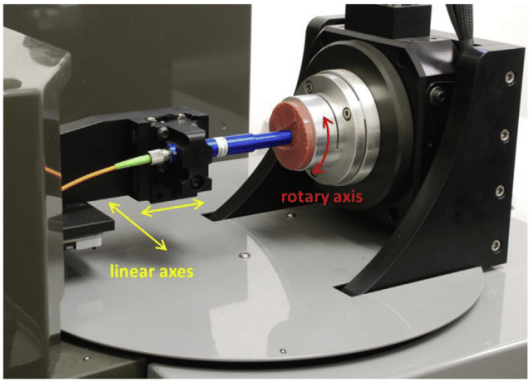


**Fig. 2.** Arrangement of the RedLux measuring device.

RedLux profiler device can be found in Ref.[48] The arrangement of the measuring device is shown in Fig. 2.

Since the measurement of a surface with a RedLux profiler is a free-form surface measurement task, it is important to scan the surface with a sufficiently dense point cloud to capture the real scanned surface with all its irregularities. Throughout all of the performed measurements, the point cloud density was 540 individual points per one rotation of the measured sample in the vertical plane; within every single rotation, the measured sample was shifted by approximately 0.1 mm in the horizontal plane. More than 120,000 points were obtained for each sample using a RedLux profiler device. The resulting point cloud for the selected sample is shown in Fig. 3.

As it is shown in Fig. 3, the point coverage on the sample surface is irregular due to the used scanning method. For a subsequent comparison with the HP 3D Structured Light Scanner Pro S2 and the NextEngine laser scanner, it was necessary to obtain a regular network of points. First, the co-ordinate data collected from the RedLux profiler were interpolated using Matlab's function *scatteredInterpolant* (The Mathworks, Inc.). This function is included in the basic Matlab package and performs interpolation on 3D sets of points that have no structure among their relative locations (scattered data set).

Interpolation step $\Delta = 0.025$ mm and a 'natural' interpolation method was chosen for both directions, specifying the density of the net. A detailed description of this function can be found in the Matlab software.[49] The mesh representation of the surface is shown in Fig. 3.

#### 3.1.2. NextEngine and HP 3D structured light scanner PRO S2

The whole sample of *ossa coxa* was digitized by two different surface scanners. We used the laser scanner NextEngine 3D Ultra HD and the HP 3D Structured Light Scanner PRO S2 (previously known as DAVID SLS 2) with a maximum resolution of 0.1 mm and 0.05 mm, respectively. These two scanners represent different scanning technologies: laser and structured light. Since the aim of this study is to compare the resulting 3D models from different surface scanners in terms of their surface representation and their possible impact on sex and age estimation, 3D models were created under the optimal conditions for the particular scanner.

The entire surface of *os coxae* was scanned to facilitate the creation of the osteological collection. The whole *ossa coxa* were used to estimate sex. However, only isolated pubic symphyseal surfaces were used for age-at-death estimation. The scanning process as well as the post-processing (aligning multiple scans of each bone together to form a final polygonal mesh) were performed in the integrated software of each scanner (ScanStudio HD and David LaserScanner v.3.10.4, respectively). In the case of HP 3D Structured Light Scanner PRO S2 (hereinafter HP 3D SLS), the calibration was done with the 120 mm pattern, which is ideal for objects of a size similar to an *os coxae*. The scanning procedure was performed against a black background and the scanned *ossa coxa* were manually rotated from both the ventral and dorsal sides to create a solid model. In the case of the NextEngine scanner, a dedicated rotational device was used to fix the bone and rotate it in front of the scanner so that each individual scan overlapped with the previous one and the next. The final scan was automatically assembled from these individual scans using the ScanStudio HD software. Each bone was attached to the device at two points: one along the iliac crest and the other just in front of the ischial tuberosity. Both points were selected to minimize the areas where the laser beam does not reach, thus, ensuring that if there is a fraction of the surface to be reconstructed, such an area is very small. These contact areas with the rotational device might have been approximately 1 $mm^2$ each.

The final 3D models were saved in the stl format. Each mesh had to be simplified (consistent simplification on 3 million faces) in order to facilitate manipulation. Subsequently, the 3D models from the NextEngine 3D scanner had to be scaled (in MeshLab software[50]) to obtain the same dimensions as the models from the HP 3D SLS. The models of the whole *ossa coxa* were used for sex estimation, while the articular surfaces of pubic symphysis that we isolated from the rest of the bone (in MeshLab software) served as input data for age-at-death estimation.

### 3.2. Comparison of resulting 3D surfaces

Once the mesh representation of the pubic symphyseal surface was created, it became important to calculate the distribution of deviations between the actual and reference surfaces. Both surfaces were overlapped using the *local best fit* feature implemented in the commercially available GOM Inspect software.[51] The function minimizes differences (square error) through all of the used points and the software tries to align the point sets so the differences are zero or as close to zero as possible. The resulting deviations were assessed with descriptive statistics (mean, standard deviation, median, and interquartile range) for all samples.

### 3.3. Biological profile assessment

#### 3.3.1. Sex estimation

Linear dimensions derived from 3D surface models taken with two

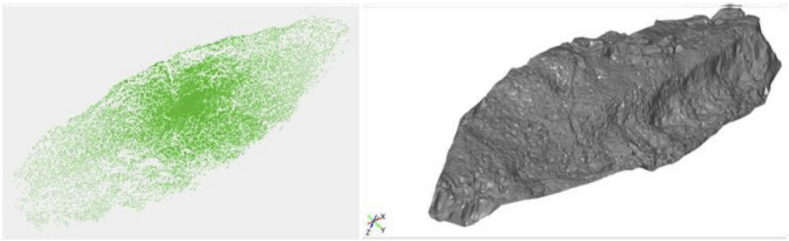

**Fig. 3.** left: Cloud of points (left), mesh representation of surface (right), sample no. 359.

**Table 1**
Os coxae measurements in DSP2.

| | Variable | Name | Reference |
|---|---|---|---|
| **1** | PUM | Acetabulo-symphyseal pubic length | [81] |
| **2** | SPU | Cotylo-pubic width | [82] |
| **3** | DCOX (M1) | Coxal length | [83] |
| **4** | IIMT (M15.1) | Greater sciatic notch height | [83] |
| **5** | ISMM | Ischium post-acetabular length | [84] |
| **6** | SCOX (12) | Iliac or coxal breadth | [83] |
| **7** | SS | Spino-sciatic length | [82] |
| **8** | SA | Spino-auricular length | [82] |
| **9** | SIS (M14.1) | Cotylo-sciatic breadth | [83] |
| **10** | VEAC (M22) | Vertical acetabular diameter | [83] |

different scanners and dimensions directly from dry *ossa coxa* served as the input data for the DSP2 (Diagnose Sexuelle Probabiliste 2) sex determination method.[52,53] The DSP2 method is based on linear discriminant analysis and posterior probabilities. The output of this sex determination tool is the probability of being male or female (sex is determined if the posterior probability is ≥ 0.95). Ten variables (or a number, at least four, that could be measured for a given bone depending on its preservation) were measured according to the definitions (Table 1) described in the software interface. The variables were first measured on dry bones with appropriate measuring instruments (sliding caliper, friction divider, and pelvimeter). Second, the measurements on the 3D models were carried out in the Morphome3cs software[54] in the virtual environment. Additionally, ten randomly chosen *ossa coxa* were measured twice over a period of four weeks (for all three types of acquisition) for intra-observer error. All the measurements were taken by two researchers (AK and RR), both trained in pelvic osteometry with approximately the same experience. Apart from the intra-observer error calculated for linear variables, the inter-observer error between the two equally trained observers was calculated as well.

To verify the accuracy of the measured dimensions within each type of data taken by one observer (intra-observer error), we calculated the technical error of measurement (TEM) and the relative TEM (*r*TEM) expressed as a percentage. Both statistical characteristics are commonly used to evaluate intra-observer precision.[27,30,55–57] The inter-observer error was computed for measurements taken on dry bones and on 3D models acquired with both scanners. A paired *t*-test was applied to assess whether there were significant differences between the two observers.

#### 3.3.2. Age-at-death estimation

The age-at-death of our sample was estimated with the use of a recent computational method of Stoyanova et al.[40,41] This fully quantitative evaluation of pubic symphyseal surfaces is based on the mathematical evaluation of the flatness of the surface, the curvature of the ventral margin of the pubic symphysis and their combination. The flatness of the surface is captured by the TPS/BE (Thin Plate Spline/Bending Energy) analysis and the SAH-Score (described in detail in Ref. 39). The last computational method captures the curvature of the ventral margin (VC) of the pubic symphysis. These analyses are incorporated into the "forAge" software. A detailed description can be found in the original study.[41]

The isolated articular surfaces of the pubic symphyses (from the reference sample as well as from both compared samples) were simplified to 15,000 faces and saved in the ply format before loading to the forAge software to estimate the age-at-death. All the isolated pubic symphyseal surfaces (n = 29) from the NextEngine, the HP 3D SLS, and the small sample (n = 5) from the RedLux were subjected to this quantitative analysis. The 3D models digitized with HP 3D SLS and NextEngine were compared using a paired *t*-test. The small sample (n = 5) of pubic bones digitized with RedLux was also compared with other samples. The differences within the corresponding individuals between RedLux and HP 3D SLS and Redlux and NextEngine, respectively, were evaluated with a paired one-sample *t*-test.

## 4. Results

### 4.1. Comparison of resulting 3D surfaces

Examples of three isolated pubic symphyseal surfaces digitized with the three scanning devices are shown in Fig. 4.

Visualization has been made in the form of color-coded maps of scatters of deviations between the reference and compared scans. This comparison was performed for all 5 selected pubic symphyseal surfaces, both for the HP 3D Structured Light Scanner Pro S2 and the NextEngine laser scanner. The color scale bar on the map of deviations shows both positive and negative dimensional changes. The positive scale (red color) means that the compared surface is above the reference surface and the negative scale (blue color) indicates that the compared surface is below the reference one. The gray color refers to the surface where no data were available from the compared scanners (HP 3D SLS and NextEngine).

An example of color-coded maps for selected surfaces is shown in Fig. 5. It can be seen that the largest and the smallest deviations from the reference surfaces are in the areas of depressions or protrusions. This is more evident on surfaces that are more rugged (e.g. pubic symphysis no. 390). It shows that scans acquired by HP 3D SLS and NextEngine are smoothed out and do not capture as much detail as the reference surface obtained with the RedLux profiler. This is due to the resolution of the particular scanning technology.

The descriptive statistics of the resulting deviations are summarized in Table 2. The graphical comparison of the examined scanners is shown in Fig. 6 for all the samples in the form of box plots. The deviations from the reference surface were slightly larger for the NextEngine than for the HP 3D SLS. In the case of the NextEngine scanner, the interquartile range is higher for all samples, as is the standard deviation, except for sample no. 383. Contrary to the NextEngine results, there are more places where no data are available in the scans from the HP 3D SLS, particularly in the major depressions. From Table 2 and Fig. 6, it can also be seen that the value of the percentage difference of *IQR* is related to the degree of the billowing of the surface. The percentage difference of *IQR* is significantly higher in the samples which are more corrugated, such as sample no. 390, where a relatively large lateral billowing is evident (Fig. 5).

### 4.2. Sex estimation

For the manual DSP method as well as for virtual DSP performed on both 3D models (HP 3D SLS and NextEngine), comparisons of estimated sexes (and agreement) between two observers were performed. The results of agreement in estimating male, female, and undetermined sex (N/A) of the two researchers are shown in Tables 3–5. Some differences between the observers were found (these variations concerned three individuals); however, the opposite sex was never assigned to the same individual (rather, such an individual was assigned N/A).

#### 4.2.1. Intra and inter-observer error of measurements

Results of intra-observer error are presented in Table 6. With the measurements on dry bones, the average TEM values were 0.80 mm and 0.62 mm, respectively. TEM values for dry bones ranged between 0.32 and 1.35 mm with *r*TEM 0.35%–3.09% for the first researcher; for the second researcher, from 0.45 to 1.01 mm, *r*TEM 0.20%–2.07%. The lowest average TEM values were reached by both researchers coincidently in the case of HP 3D SLS (0.62 and 0.56 mm, respectively). The intra-observer error of the Researcher 1 ranged between 0.35 and 1.24 mm for all ten dimensions and the *r*TEM ranged between 0.21% and 2.85%; for Researcher 2, between 0.24 and 0.88 mm with *r*TEM

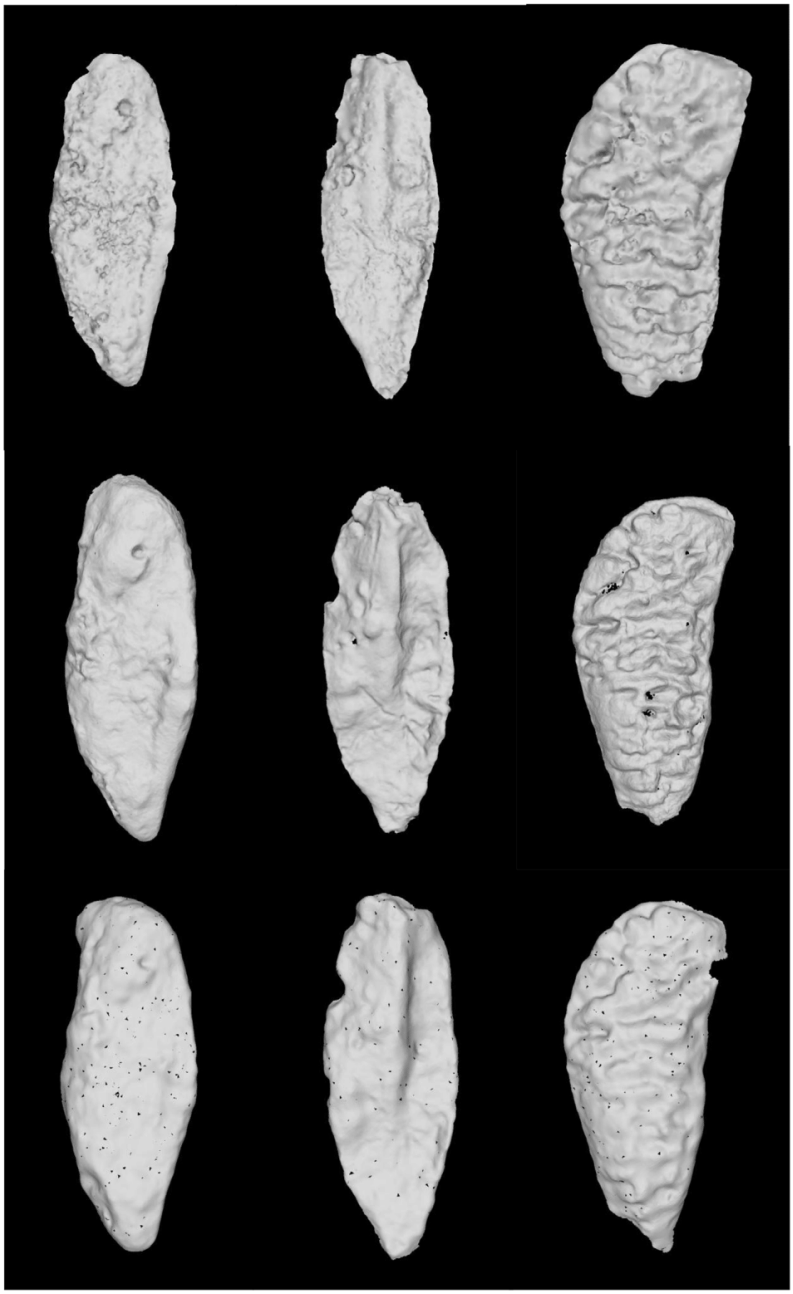

**Fig. 4.** Examples of digitized surfaces of isolated pubic symphyses. Models derived from RedLux (upper row), HP 3D SLS (middle row) and NextEngine (lower row).

from 0.19% to 2.22%. The average intra-observer error for NextEngine was very similar to the previous results (0.72 mm and 0.57 mm, respectively). The error for Researcher 1 was between 0.26 and 1.38 mm, *r*TEM 0.20%–2.36%; the TEM of Researcher 2 ranged from 0.12 to 1.19 mm, with *r*TEM of 0.07%–1.95%.

The results of inter-observer error for the variables measured directly on dry bones, on 3D models digitized with HP 3D SLS, and NextEngine, respectively, are shown in Table 7. Significant differences between the two researchers when dry bones were measured were found for IIMT, ISMM, SCOX, SS, and VEAC). The DSP variables measured on the 3D models from the HP 3D SLS scanner were statistically different between the two researchers in IIMT, ISMM, SCOX, SS, SIS, and VEAC. Inter-observer differences were also found for the 3D measurements derived from NextEngine, namely for PUM, DCOX, SS, and VEAC.

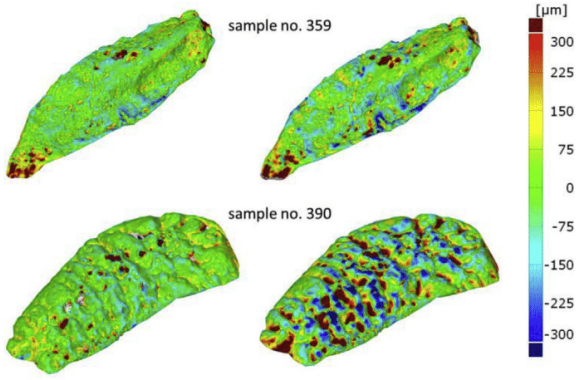


**Fig. 5.** Color-coded maps of the resulting deviations (sample no. 359 and 390); left: HP 3D SLS vs. reference surface; right: NextEngine laser scanner vs. reference surface. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

### 4.3. Age estimation

A paired *t*-test was applied to reveal the potential differences between the estimated ages derived from the 3D models of the isolated pubic symphyses from the NextEngine and the HP 3D SLS. The results are provided in Table 8. The differences in the estimated ages were not statistically significant. The only exception was the TPS/BE regression model. Thus, the null hypothesis was not rejected for any of the models, except for the TPS/BE model. These results can be interpreted as a negligible effect of the scanning technologies on age estimation in this study. The sample of isolated pubic symphyseal surfaces derived from the RedLux Profiler (n = 5) was also subjected to the age-at-death analysis. We did detect differences between HP 3D SLS and RedLux for the SAH model, and between NextEngine and Redlux in the SAH and SAH + VC models. The null hypothesis was not rejected in all other cases (see Table 9).

## 5. Discussion

### 5.1. Scanning – problem of compatibility

Many researchers travel around the world and study osteological collections to design new methods or to validate already existing ones. Scanning technologies are already used in biological and forensic anthropology and have become of paramount importance (e.g. Refs. 4,7,58). Some of the greatest benefits are the digital storage and sharing of osteological remains. The digital material can come to researchers or it can be publicly available under certain conditions. However, research institutions are equipped with different scanning devices that may produce varying outputs.

Therefore, it is necessary to verify that the outputs from different devices are comparable to each other (and under what settings) and that they do not affect further analyses such as the estimation of the biological profile. The comparability of outputs from different devices is especially important in age-at-death estimation since minor changes on the joint surface may influence the final estimate significantly. However, the consistency of outputs from different devices for biological profile estimation needs to be tested. This has not yet been sufficiently explored, except for a few exceptions, such as the aforementioned study of Villa et al.[43]

To test the effect of the scanning device on the forensic methods of biological profile estimation, we chose sex and age-at-death estimation. For sexing, the method called DSP is well established; its reliability has been proven and it guarantees objectivity (e.g. Refs. 27,28,59,60). The forAge software is one of the few age estimation approaches using the mathematical quantification of bone surface that has been published so far. The repeatability of the method was tested with excellent results[61] and its use is suitable for individuals under 40 years of age.[62]

The advantages and disadvantages of both scanners (NextEngine and HP 3D Structured Light Scanner PRO S2) and the technologies represented by them should be discussed to help future users assess them for their research goals. Both scanning technologies are used on skeletal remains. However, laser scanning technology is generally preferred among anthropologists.[18,63] One of the most often used scanners is the NextEngine laser scanner (e.g. Refs. 7,10,39,64–66), which was also used for the development of one method of age estimation that we tested in our study.[40,41] Although less attention has been paid to HP 3D SLS, its use is now on the rise. HP 3D SLS was used or tested in many studies on human and animal bones, as well as for other purposes.[19,67–74] Both tested scanners are commercially available, and belong to the low-cost category of scanners (under $ 3,000), which makes them accessible to a great number of users. Moreover, the cost vs. performance ratio very often makes them the first choice. In terms of the time needed to acquire and post-process scans, the HP 3D SLS outperforms the NextEngine scanner, mainly because the laser technology is more time-consuming than other technologies.[4,15] Both scanners are characterized by high portability, and can be easily carried to various collections around the world. Also, both scanners are able to capture texture and provide a 3D textured mesh as an output, which is often a great benefit.[7]

Nonetheless, both NextEngine and HP 3D SLS have some limitations. Although laser and white light scanners are commonly used on skeletal remains, they are limited in scanning dark or transparent objects, or objects reflecting light (e.g. tooth enamel).[7,75,76] Systems using blue light (e.g. HP 3D SLS) should reduce this limitation[7]; however, one of the authors (A.K.), who routinely operates the HP 3D SLS, has experienced similar issues (problems capturing very dark areas on objects). In the case of the HP 3D SLS, constant ambient light during scanning is recommended, and a generally darker room is better suited for scanning.

**Table 2**
Results of quantitative comparison between measured samples (HP 3D SLS vs. Reference sample[a] and NextEngine vs. Reference sample).

| No. of individual | Scanner | Mean value for normal distribution | Median | Standard deviation | Interquartile span | Percentage difference *IQR* |
|---|---|---|---|---|---|---|
| 352 | HP 3D | 26 | 2 | 154 | 147 | +50 |
| | NextE | 26 | 4 | 204 | 220 | |
| 359 | HP 3D | 8 | −8 | 118 | 104 | +35 |
| | NextE | 7 | −4 | 146 | 140 | |
| 382 | HP 3D | 14 | 6 | 113 | 114 | +24 |
| | NextE | 13 | 0 | 138 | 141 | |
| 383 | HP 3D | 15 | 5 | 187 | 122 | +51 |
| | NextE | 26 | 15 | 174 | 184 | |
| 390 | HP 3D | 14 | 0 | 126 | 88 | +141 |
| | NextE | 21 | 2 | 202 | 212 | |

[a] Reference samples were acquired with the RedLux Profiler.

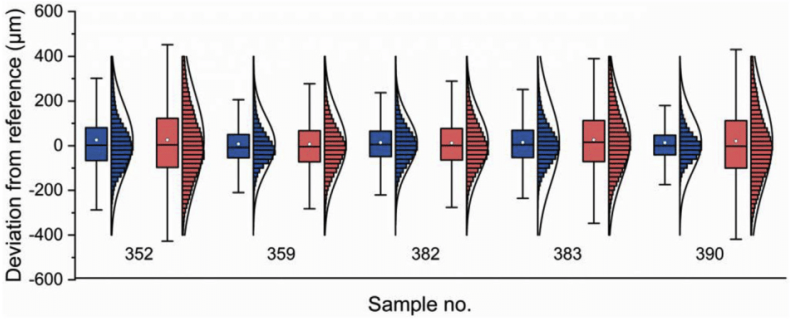


**Fig. 6.** Deviation from the reference sample (Redlux scan) for the HP 3D SLS (blue) and NextEngine laser scanner (red). On each box, the central mark is the median, the white point is the mean value for normal distribution $\mu$, the edges of the box are the 25th and 75th percentiles, and the whiskers extend to the most extreme data points not considered outliers. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

**Table 3**
Agreement between researchers in estimating sex with the use of DSP on dry bones. Shaded cells correspond to the number of individuals in which the two researchers agreed.

| | | Researcher 1 | | | |
|---|---|---|---|---|---|
| | | Male | Female | N/A | Total |
| Researcher 2 | Male | 15 | 0 | 0 | 15 |
| | Female | 0 | 13 | 0 | 13 |
| | N/A | 1 | 0 | 0 | 1 |
| | Total | 16 | 13 | 0 | 29 |

**Table 4**
Agreement between the researchers in estimating sex with the use of the DSP on virtual models from HP 3D SLS. Shaded cells correspond to the number of individuals in which the two researchers agreed.

| | | Researcher 1 | | | |
|---|---|---|---|---|---|
| | | Male | Female | N/A | Total |
| Researcher 2 | Male | 15 | 1 | 0 | 16 |
| | Female | 0 | 11 | 0 | 11 |
| | N/A | 0 | 1 | 1 | 2 |
| | Total | 15 | 13 | 1 | 29 |

**Table 5**
Agreement between the two researchers in estimating sex with the use of the DSP on virtual models from NextEngine. Shaded cells correspond to the number of individuals in which the two researchers agreed.

| | | Researcher 1 | | | |
|---|---|---|---|---|---|
| | | Male | Female | N/A | Total |
| Researcher 2 | Male | 15 | 0 | 0 | 15 |
| | Female | 0 | 12 | 0 | 12 |
| | N/A | 1 | 0 | 1 | 2 |
| | Total | 16 | 12 | 1 | 29 |

It has to be emphasized that the quality of 3D models could be affected by post-processing. Simplification is often required to reduce file size and to fasten data processing and manipulation to virtual data.[4,77] Even though in the case of smoothing procedures the positive or negative effect on mesh topology depends on the algorithm used, the decimation procedure always leads to loss of information, which affects mesh topology and measurement. By increasing the mesh triangle size, the accuracy of measurement is negatively affected.[77,78] The amount of decimation should be defined with regards to the type of object and the purpose for which it is to be used.[4,77] In this paper, the only post-processing step performed was decimation, in order to enable better manipulation on the computer during analysis. Since the whole *ossa coxa* was scanned, in the case of HP 3D SLS and NextEngine, simplification was applied to all the models equally to enhance the possibility of manipulation with minimum information loss. These models were then decimated once more due to the forAge software properties. The reference models using Redlux needed to be simplified only once in order to estimate age. Other post-processing procedures (e.g. hole filling, smoothing) were not performed to prevent the quality of the scan from being negatively affected. Therefore, the dimensions in our study could not be affected by different post-processing procedures. However, it would be of great benefit for future studies if somebody tested to what extent the surface could be simplified without losing too much information or obtaining inaccurate or imprecise estimates.

In our study, we worked with the assumption that different scanner operators or different scanning protocols do not affect the resulting 3D models or their geometric properties.[78] However, all the virtual data were acquired by experienced operators to ensure the consistency of scanned data. All the data were taken under the optimal conditions of each scanner. Since it would be complicated to compare scanning devices with each other, such an approach seemed the most reasonable to us. Moreover, it was also used in the study by Villa et al.[43]

### 5.2. Comparison of tested devices

Even the visual inspection of the final 3D models revealed that the Redlux scanner captures much more detail than the other two low-cost scanners.

Our results indicate that the NextEngine laser scanner captured fewer details (it smoothed the surface the most and had the largest information loss) than the HP 3D Structured Light Scanner when compared to the reference scans from the Redlux Profiler. Singh et al.[79] report that structured light technology is more accurate in comparison to laser technology when used to calculate surface area and volume. Results of another study[15] indicate that structured light technology captured external structures better than other technologies (including laser technology). It should be noted here that these differences, as found in our study, are relatively small and their impact on further analyses (biological profile analyses) would be negligible. Our analyses of sex and age estimation confirm this notion.

### 5.3. Influence of different scanning devices on the linear measurement and sex assessment

Both observers estimated the sex of all individuals almost uniformly irrespective of the scanning technology. The intra-observer and inter-observer errors of the linear measurements were evaluated. According to Camison et al.,[80] who divided *r*TEM values into five categories (< 1% = excellent, 1–3.9% = very good, 4–6.9% = good, 7–9.9% = moderate, and > 10% = poor), the intra-observer error of both researchers in most cases fell into the "excellent" category and the rest into "very good" (the highest value reached 3.09%). Nevertheless, our results of inter-observer error show some statistically significant differences between the two observers, which may not be due to different scanners, as their results never led to the opposite classification.

**Table 6**
The intra-observer precision of measured variables used for DSP.

| | Researcher 1 | | | | | | Researcher 2 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | dry bones | | HP 3D SLS | | NextEngine | | dry bones | | HP 3D SLS | | NextEngine | |
| Variable | TEM | *r*TEM | TEM | *r*TEM | TEM | *r*TEM | TEM | *r*TEM | TEM | *r*TEM | TEM | *r*TEM |
| PUM | 1.19 | 1.58 | 1.24 | 1.64 | 1.38 | 1.84 | 0.56 | 0.74 | 0.82 | 1.07 | 1.17 | 1.51 |
| SPU | 0.89 | 3.09 | 0.80 | 2.85 | 0.67 | 2.36 | 0.56 | 1.97 | 0.63 | 2.22 | 0.46 | 1.57 |
| DCOX | 0.77 | 0.35 | 0.49 | 0.23 | 1.11 | 0.51 | 0.45 | 0.20 | 0.83 | 0.38 | 1.19 | 0.55 |
| IIMT | 1.35 | 2.91 | 0.91 | 2.05 | 0.78 | 1.76 | 0.94 | 2.07 | 0.88 | 2.02 | 0.85 | 1.95 |
| ISMM | 1.25 | 1.14 | 0.57 | 0.52 | 1.31 | 1.16 | 0.47 | 0.43 | 0.67 | 0.60 | 0.55 | 0.49 |
| SCOX | 0.57 | 0.35 | 0.35 | 0.21 | 0.33 | 0.20 | 0.55 | 0.34 | 0.30 | 0.19 | 0.12 | 0.07 |
| SS | 0.43 | 0.57 | 0.43 | 0.58 | 0.26 | 0.35 | 0.60 | 0.81 | 0.24 | 0.33 | 0.47 | 0.63 |
| SA | 0.68 | 0.84 | 0.53 | 0.67 | 0.40 | 0.50 | 1.01 | 1.25 | 0.38 | 0.48 | 0.42 | 0.52 |
| SIS | 0.32 | 0.80 | 0.40 | 1.03 | 0.38 | 0.97 | 0.49 | 1.24 | 0.43 | 1.11 | 0.26 | 0.66 |
| VEAC | 0.54 | 0.95 | 0.52 | 0.91 | 0.58 | 1.00 | 0.60 | 1.07 | 0.45 | 0.81 | 0.24 | 0.43 |
| mean | 0.80 | | 0.62 | | 0.72 | | 0.62 | | 0.56 | | 0.57 | |

TEM = technical error of measurement; *r*TEM = relative technical error of measurement; the measurement unit of TEM is given in mm; *r*TEM expressed as a percentage.

**Table 7**
Results of measurement differences (used for sex estimation with the DSP) on dry bones, 3D models made with HP 3D SLS and NextEngine between two researchers.

| Variable | N | Max difference (mm) | Mean difference (mm) | SD | p-value |
|---|---|---|---|---|---|
| **Dry bones** | | | | | |
| PUM | 29 | 4.00 | 1.22 | 4.99 | 0.05 |
| SPU | 28 | 4.00 | 1.34 | 3.06 | 0.11 |
| DCOX | 27 | 4.00 | 0.98 | 9.18 | 0.53 |
| IIMT | 28 | 4.50 | 1.79 | 6.79 | < 0.005 |
| ISMM | 29 | 3.50 | 1.07 | 5.98 | < 0.001 |
| SCOX | 15 | 3.00 | 0.83 | 6.27 | 0.01 |
| SS | 28 | 3.50 | 1.70 | 3.64 | < 0.001 |
| SA | 28 | 4.50 | 1.70 | 4.34 | 0.12 |
| SIS | 29 | 2.00 | 0.45 | 2.9 | 0.27 |
| VEAC | 29 | 3.50 | 0.97 | 2.99 | 0.02 |
| **HP 3D SLS** | | | | | |
| PUM | 29 | 5.00 | 1.66 | 4.88 | 0.18 |
| SPU | 28 | 3.30 | 1.04 | 3.01 | 0.22 |
| DCOX | 27 | 6.49 | 2.18 | 8.49 | 0.21 |
| IIMT | 28 | 5.24 | 1.79 | 6.86 | 0.02 |
| ISMM | 29 | 5.00 | 1.01 | 5.99 | < 0.01 |
| SCOX | 15 | 3.81 | 1.11 | 6.49 | 0.04 |
| SS | 28 | 3.32 | 0.96 | 3.66 | < 0.001 |
| SA | 28 | 4.10 | 0.97 | 4.61 | 0.89 |
| SIS | 29 | 1.01 | 0.51 | 2.93 | < 0.01 |
| VEAC | 29 | 4.97 | 1.90 | 3.07 | < 0.001 |
| **NextEngine** | | | | | |
| PUM | 29 | 7.76 | 2.51 | 4.63 | < 0.001 |
| SPU | 28 | 3.49 | 1.16 | 3.14 | 0.07 |
| DCOX | 27 | 7.46 | 2.13 | 9.25 | 0.01 |
| IIMT | 28 | 4.88 | 1.49 | 7.08 | 0.22 |
| ISMM | 29 | 2.60 | 0.82 | 6.21 | 0.83 |
| SCOX | 15 | 5.18 | 0.97 | 6.60 | 0.20 |
| SS | 28 | 2.85 | 1.19 | 3.67 | < 0.001 |
| SA | 28 | 3.01 | 0.98 | 4.68 | 0.51 |
| SIS | 29 | 2.09 | 0.54 | 2.75 | 0.44 |
| VEAC | 29 | 4.31 | 2.39 | 3.32 | < 0.001 |

Here, our results could be interpreted in a way that is similar to the study of Mullins and Albanesse[32]: that even significant differences (in the observer) need not affect the estimates or utility of this type of data. To the best of our knowledge, this is the first study testing the DSP method on virtual models acquired by surface scanners. Previously, the method was tested only on 3D models acquired with CT scanning (e.g. Refs. 27,28,60).

**Table 8**
Results of a paired *t*-test between the estimated ages based on models from HP 3D SLS and NextEngine.

| Regression model | p-value |
|---|---|
| TPS/BE[a] | 0.01 |
| SAH[b] | 0.07 |
| VC[c] | 0.10 |
| TPS/BE + VC[d] | 0.27 |
| SAH + VC[e] | 0.41 |

P-values are shown for a two-tailed test.
[a] Thin plate spline/Bending energy.
[b] Slice and Algee-Hewitt score.
[c] Ventral curvature.
[d] Combination of TPS/BE and VC (multivariate model).
[e] Combination of SAH score and VC (multivariate model).

**Table 9**
Results of a paired *t*-test for differences between estimated ages derived from the HP 3D SLS vs. RedLux and between NextEngine vs RedLux.

| | HP 3D SLS vs. RedLux | NextEngine vs. Redlux |
|---|---|---|
| Regression model | p-value | |
| TPS/BE[a] | 0.74 | 0.41 |
| SAH[b] | 0.04 | 0.03 |
| VC[c] | 0.30 | 0.18 |
| TPS/BE + VC[d] | 0.87 | 0.99 |
| SAH + VC[e] | 0.05 | 0.03 |

P-values are shown for a two-tailed test.
[a] Thin plate spline/Bending energy.
[b] Slice and Algee-Hewitt score.
[c] Ventral curvature.
[d] Combination of TPS/BE and VC (multivariate model).
[e] Combination of SAH score and VC (multivariate model).

### *5.4. Effect of the scanning device on delicate surface changes in the age-at-death estimation*

Although the comparison of captured surfaces from both low-cost scanners showed slightly better outputs for the HP 3D SLS, there was almost no effect detected on the quantitative methods of age estimation (the only exception was the TPS/BE model). We can conclude that other scanning technologies (at least structured light, here represented by the HP 3D SLS) can be used to estimate age using the method of Stoyanova

et al.[40,41]

When compared to the outputs obtained with the Redlux, significant differences in one out of five models were found in the case of HP 3D SLS and two out of five with NextEngine. However, it was not assumed in our study that Redlux should be used to estimate age-at-death through quantitative methods. This device was selected only as a reference for comparing two low cost scanners as its routine use in anthropology would be too expensive and impractical.

Larger deviations between the captured surfaces are more evident on more billowed surfaces, e.g. sample no. 390, which was obviously a very young individual (with a partially unfused iliac crest). Therefore, further investigation is needed to clarify whether the effect of different scanners is more obvious in more billowed surfaces, typical for younger individuals.

## 6. Conclusion

The two tested scanning technologies, structured light and laser represented by HP 3D SLS and NextEngine, respectively, showed only small surface deviation from the reference sample in our study. The structured light technology seems to be more accurate (captures slightly more detail) than the laser technology. Nevertheless, no significant impact on age and sex estimates was observed, except for the TPS/BE model in age estimation. Thus, it appears that the type of scanner does not have a significant effect on the estimate of the biological profile under optimal scanning settings. However, we encourage further investigation, especially in the case of age-at-death estimation, where even subtle changes to the articulation surface are evaluated. Both the existing as well as future analyses of these surfaces could be sensitive to the different 3D representations acquired with various devices.

## Declarations of interest

None.

## Acknowledgments

We would like to thank the Department of Anthropology at the National Museum in Prague for providing the osteological material.

This research has been supported by the research grant Charles University Grant Agency No. 642218 and partially by IRN Bipedal Equilibrium, CNRS, France.

## References

1. Villa C, Buckberry J, Lynnerup N. Evaluating osteological ageing from digital data. *J Anat.* 2016https://doi:10.1111/joa.12544.
2. Friess M. Calvarial shape variation among Middle Pleistocene hominins: an application of surface scanning in paleoanthropology. *CR Palevol.* 2010;9:435–443https://doi:10.1016/j.crpv.2010.07.016.
3. Kuzminsky SC, Gardiner MS. Three-dimensional laser scanning : potential uses for museum conservation and scientific research. *J Archaeol Sci.* 2012;39:2744–2751https://doi:10.1016/j.jas.2012.04.020.
4. Friess M. Scratching the surface? The use of surface scanning in physical and paleoanthropology. *J Anthropol Sci.* 2012;90:1–25https://doi:10.4436/jass.90004.
5. Thompson TJU, Norris P. A new method for the recovery and evidential comparison of footwear impressions using 3D structured light scanning. *Sci Justice.* 2018;58:237–243https://doi:10.1016/j.scijus.2018.02.001.
6. Naether S, Buck U, Raess B, Thali M. Crime scene reconstruction using 3-D scanning and medical imaging technologies. *Sci Justice.* 2010;50:35https://doi:10.1016/j.scijus.2009.11.037.
7. Slizewski A, Friess M, Semal P. Surface scanning of anthropological specimens: nominal-actual comparison with low cost laser scanner and high end fringe light projection surface scanning systems. *Quartar.* 2010;57:179–187.
8. Viggiano D, Thanassoulas T, Di-Cesare C, et al. A low-cost system to acquire 3D surface data from anatomical samples. *Eur J Anat.* 2015;19:343–349.
9. Aung SC, Ngim RCK, Lee ST. Evaluation of the laser scanner as a surface measuring tool and its accuracy compared with direct facial anthropometric measurements. *Br J Plast Surg.* 1995;48:551–558https://doi:10.1016/0007-1226(95)90043-8.
10. Cantín M, Muñoz M, Olate S. Generation of 3D tooth models based on three-dimensional scanning to study the morphology of permanent teeth. *Int J Morphol.* 2015;33:782–787.
11. Gibelli D, Pucciarelli V, Poppa P, et al. Three-dimensional facial anatomy evaluation: reliability of laser scanner consecutive scans procedure in comparison with stereophotogrammetry. *J Cranio-Maxillo-Fac Surg.* 2018;46:1807–1813https://doi:10.1016/j.jcms.2018.07.008.
12. Adams JW, Olah A, McCurry MR, Potze S, Wilson BA. Surface model and tomographic archive of fossil primate and other mammal holotype and paratype specimens of the Ditsong National Museum of Natural History, Pretoria, South Africa. *PLoS One.* 2015;10:1–14https://doi.org/10.1371/journal.pone.0139800.
13. Mcpherron SP, Gernat T, Hublin J. Structured light scanning for high-resolution documentation of in situ archaeological finds. *J Archaeol Sci.* 2009;36:19–24https://doi:10.1016/j.jas.2008.06.028.
14. Guydish M, Henson K. Using digitized Native American skeletal remains to conduct osteological analyses. *Proc W Va Acad Sci.* 2017;89.
15. Mathys A, Brecko J, Semal P. Comparing 3D digitizing technologies: what are the differences? *Digital Heritage International Congress (DigitalHeritage).* IEEE; 2013:201–204.
16. Errickson D. Shedding light on skeletal remains: the use of structured light scanning for 3D archiving. In: Errickson D, Thompson T, eds. *Human Remains: Another Dimension.* Academic Press; 2017:93–101.
17. Ferreira MT, Ross AH, Cunha E. A reflection on the maintenance of identified skeletal collections state of preservation. *Rev Med Leg.* 2017;8:186https://doi.org/10.1016/j.medleg.2017.10.017.
18. Errickson D, Grueso I, Griffith SJ, et al. Towards a best practice for the use of active non-contact surface scanning to record human skeletal remains from archaeological contexts. *Int J Osteoarchaeol.* 2017;27:650–661https://doi:10.1002/oa.2587.
19. Klein S, Avery M, Adams G, Pollard S, Simske S. From scan to print: 3D printing as a means for replication. *NIP & Digital Fabrication Conference. Society for Imagining Science and Technology.* 2014; 2014:417–421.
20. Carew RM, Morgan RM, Rando C. A preliminary investigation into the accuracy of 3D modeling and 3D printing in forensic anthropology evidence reconstruction. *J Forensic Sci.* 2018;64:342–352https://doi:10.1111/1556-4029.13917.
21. Chase RJ, LaPorte G. The next generation of crime tools and challenges: 3D printing. *Natl Inst Justice.* 2017;279:49–57.
22. Rmoutilová R, Guyomarc'h P, Velemínský P, et al. Virtual reconstruction of the upper palaeolithic skull from Zlatý Kůň, Czech republic: sex assessment and morphological affinity. *PLoS One.* 2018;13:e0201431.
23. Weber G, Schäfer K, Prossinger H, Gunz P, Mittereocker P, Seidler H. Virtual anthropology: the digital evolution in anthropological sciences. *J Physiol Anthropol Appl Hum Sci.* 2001;20:69–80.
24. Benazzi S, Gruppioni G, Strait DS, Hublin JJ. Technical Note: virtual reconstruction of KNM-ER 1813 Homo habilis cranium. *Am J Phys Anthropol.* 2014;153:154–160https://doi:10.1002/ajpa.22376.
25. Langley-Shirley N, Tersigni-Tarrant MA. *Forensic Anthropology A Comprehensive Introduction.* second ed. Boca Raton: CRC Press; 2017.
26. Nikita E. *Osteoarchaeology: A Guide to the Macroscopic Study of Human Skeletal Remains.* London: Academic Press; 2017.
27. Mestekova S, Bruzek J, Veleminska J, Chaumoitre K. A test of the DSP sexing method on CT images from a modern French sample. *J Forensic Sci.* 2015;60:1295–1299https://doi:10.1111/1556-4029.12817.
28. Chapman T, Lefevre P, Semal P, et al. Sex determination using the probabilistic sex diagnosis (DSP : Diagnose sexuelle Probabiliste) tool in a virtual environment. *Forensic Sci Int.* 2014;234 189-e1 https://doi:10.1016/j.forsciint.2013.10.037.
29. Citardi MJ, Herrmann B, Hollenbeak CS, Stack BC, Cooper M, Bucholz RD. Comparison of scientific calipers and computer-enabled CT review for the measurement of skull base and craniomaxillofacial dimensions. *Skull Base.* 2001;11:5–11https://doi:10.1055/s-2001-12781.
30. Corron L, Marchal F, Condemi S, Chaumoître K, Adalian P. Evaluating the consistency, repeatability, and reproducibility of osteometric data on dry bone surfaces, scanned dry bone surfaces, and scanned bone surfaces obtained from living individuals. *BMSAP.* 2017;29:33–53https://doi:10.1007/s13219-016-0172-7.
31. Verhoff MA, Ramsthaler F, Krähahn J, et al. Digital forensic osteology-Possibilities in cooperation with the Virtopsy®project. *Forensic Sci Int.* 2008;174:152–156https://doi:10.1016/j.forsciint.2007.03.017.
32. Mullins RA, Albanese J. Estimating biological characteristics with virtual laser data. *J Forensic Sci.* 2018;63:815–823https://doi:10.1111/1556-4029.13621.
33. Musilová B, Dupej J, Velemínská J, Chaumoitre K. Exocranial surfaces for sex assessment of the human cranium. *Forensic Sci Int.* 2016;269:70–77https://doi:10.1016/j.forsciint.2016.11.006.
34. Abdel Fatah EE, Shirley NR, Jantz RL, Mahfouz MR. Improving sex estimation from crania using a novel three-dimensional quantitative method. *J Forensic Sci.* 2014;59:590–600https://doi:10.1111/1556-4029.12379.
35. Bulut O, Petaros A, Hizliol I, Wärmländer SKTS, Hekimoglu B. Sexual dimorphism in frontal bone roundness quantified by a novel 3D-based and landmark-free method. *Forensic Sci Int.* 2016;261 162.e1-162.e5 https://doi:10.1016/j.forsciint.2016.01.028.
36. Cavaignac E, Li K, Faruch M, et al. Three-dimensional geometric morphometric analysis reveals ethnic dimorphism in the shape of the femur. *J Exp Othopaedics.* 2017;4:13https://doi:10.1186/s40634-017-0088-2.
37. Murphy RE, Garvin HM. A morphometric outline analysis of ancestry and sex differences in cranial shape. *J Forensic Sci.* 2018;63:1001–1009https://doi:10.1111/1556-4029.13699.
38. Villa C, Buckberry J, Cattaneo C, Frohlich B, Lynnerup N. Quantitative analysis of the morphological changes of the pubic symphyseal face and the auricular surface and implications for age at death estimation. *J Forensic Sci.* 2015;60:556–565https://doi.org/10.1111/1556-4029.12689.
39. Slice DE, Algee-Hewitt BFB. Modeling bone surface morphology: a fully quantitative

method for age-at-death estimation using the pubic symphysis. *J Forensic Sci.* 2015;60:835–843https://doi.org/10.1111/1556-4029.12778.
40. Stoyanova D, Algee-Hewitt BFB, Slice DE. An enhanced computational method for age-at-death estimation based on the pubic symphysis using 3D laser scans and thin plate splines. *Am J Phys Anthropol.* 2015;158:431–440https://doi.org/10.1002/ajpa.22797.
41. Stoyanova DK, Algee-Hewitt BFB, Kim J, Slice DE. A computational framework for age-at-death estimation from the skeleton: surface and outline analysis of 3D laser scans of the adult pubic symphysis. *J Forensic Sci.* 2017;62:1434–1444https://doi.org/10.1111/1556-4029.13439.
42. San-Millán M, Rissech C, Turbón D. Shape variability of the adult human acetabulum and acetabular fossa related to sex and age by geometric morphometrics. Implications for adult age estimation. *Forensic Sci Int.* 2017;272:50–63https://doi.org/10.1016/j.forsciint.2017.01.005.
43. Villa C, Gaudio D, Cattaneo C, Buckberry J, Wilson AS, Lynnerup N. Surface curvature of pelvic joints from three laser scanners: separating anatomy from measurement error. *J Forensic Sci.* 2015;60:374–381https://doi.org/10.1111/1556-4029.12696.
44. Poláček L. Das Hinterland des frühmitelalterlichen Zentrums in Mikulčice. Stand und Perspektiven der Forschung. In: Poláček L, ed. *Das Wirtschäftliche Hinterl. Der Frühmittelalterlichen Zentren. Int. Tagungen Mikulčice VI, Archeologický Ústav Akademie Věd ČR.* 2008; 2008:257–298 Brno.
45. Nawabi DH, Nassif NA, Do HT, et al. What causes unexplained pain in patients with metal-on metal hip devices? A retrieval, histologic, and imaging analysis. *Clin Orthop Relat Res.* 2014;472:543–554https://doi:10.1007/s11999-013-3199-9.
46. Cook RB, Shearwood-Porter NR, Latham JM, Wood RJK. Volumetric assessment of material loss from retrieved cemented metal hip replacement stems. *Tribiology Int.* 2015;89:105–108https://doi:10.1016/j.triboint.2014.12.026.
47. ACC SILICONES LTD. *Technical Data Sheet.* 2017; 2017https://acc-silicones.com/products/moulding_rubbers/MM242R.
48. Tuke M, Taylor A, Roques A, Maul C. 3D linear and volumetric wear measurement on artificial hip joints-Validation of a new methodology. *Precis Eng.* 2010;34:777–783https://doi.org/10.1016/j.precisioneng.2010.06.001.
49. The MathWorks, Inc. *MATLAB R2015b [software].* 2015; 2015.
50. Cignoni P, Callieri M, Corsini M, Dellepiane M, Ganovelli F, Ranzuglia G. *MeshLab: An Open-Source Mesh Processing Tool, 6th Eurographics Italian Chapter Conference.* 2008; 2008:129–136.
51. GOM GmbH. *GOM Inspect 2016 [software].* 2016; 2016.
52. Murail P, Bruzek J, Houët F, Cunha E. DSP: a tool for probabilistic sex diagnosis using worldwide variability in hip-bone measurements. *BMSAP.* 2005;17:167–176.
53. Brůžek J, Santos F, Dutailly B, Murail P, Cunha E. Validation and reliability of the sex estimation of the human os coxae using freely available DSP2 software for bioarchaeology and forensic anthropology. *Am J Phys Anthropol.* 2017;164:440–449https://doi:10.1002/ajpa.23282.
54. Morphome3cs II. *CGG MFF UK.* 2015; 2015http://www.morphome3cs.com/.
55. Lottering N, Reynolds MS, MacGregor DM, Meredith M, Gregory LS. Morphometric modelling of ageing in the human pubic symphysis: sexual dimorphism in an Australian population. *Forensic Sci Int.* 2014;236 195.e1-e11 https://doi.org/10.1016/j.forsciint.2013.12.041.
56. Stomfai S, Ahrens W, Bammann K, et al. Intra- and inter-observer reliability in anthropometric measurements in children. *Int J Obes.* 2011;35:45–51https://doi:10.1038/ijo.2011.34.
57. Machado MPS, Costa ST, Freire AR, et al. Application and validation of Diagnose Sexuelle Probabiliste V2 tool in a miscegenated population. *Forensic Sci Int.* 2018;290 351.e1-351.e5 https://doi:10.1016/j.forsciint.2018.06.043.
58. Carew RM, Errickson D. Imaging in forensic science : five years on. *J For Radiol Imag.* 2019;16:24–33https://doi:10.1016/j.jofri.2019.01.002.
59. Quatrehomme G, Radoman I, Nogueira L, du Jardin P, Alunni V. Sex determination using the DSP (probabilistic sex diagnosis) method on the coxal bone: efficiency of method according to number of available variables. *Forensic Sci Int.* 2017;272:190–193https://doi:10.1016/j.forsciint.2016.10.020.
60. Rodríguez Paz A, Banner J, Villa C. Validity of the probabilistic sex diagnosis method (DSP) on 3D CT-scans from modern Danish population. *Rev Med Leg.* 2018https://doi:10.1016/j.medleg.2018.08.002.
61. Kim J, Algee-Hewitt BFB, Stoyanova DK, Figueroa-Soto C, Slice D. Testing reliability of the computational age-at-death estimation methods between five observers using three-dimensional image data of the pubic symphysis. *J Forensic Sci.* 2018;64:507–518https://doi:10.1111/1556-4029.13842.
62. Kotěrová A, Velemínská J, Cunha E, Brůžek J. A validation study of the Stoyanova et al. method (2017) for age-at-death estimation quantifying the 3D pubic symphyseal surface of adult males of European populations. *Int J Leg Med.* 2018;133:603–612https://doi:10.1007/s00414-018-1934-1.
63. Errickson D, Thompson T, Rankin B. *An Optimum Guide for the Reduction of Noise Using a Surface Scanner for Digitising Human Osteological Remains.* 2015; 2015https://guides.archaeologydataservice.ac.uk/g2gp/CS_StructuredLight.
64. Algee-Hewitt BFB, Wheat AD. The reality of virtual anthropology: comparing digitizer and laser scan data collection methods for the quantitative assessment of the cranium. *Am J Phys Anthropol.* 2016;160:148–155https://doi:10.1002/ajpa.22932.
65. Gualdi-Russo E, Zaccagni L, Russo V. Giovanni Battista Morgagni: facial reconstruction by virtual anthropology. *Forensic Sci Med Pathol.* 2015;11:222–227https://doi:10.1007/s12024-015-9665-9.
66. Shearer BM, Sholts SB, Garvin HM, Wärmländer SKTS. Sexual dimorphism in human browridge volume measured from 3D models of dry crania: a new digital morphometrics approach. *Forensic Sci Int.* 2012;222 e1–400. e5 https://doi:10.1016/j.forsciint.2012.06.013.
67. Tambusso PS, Mcdonald HG, Fariña RA. Description of the stylohyal bone of a giant sloth (Lestodon armatus). *Palaeontol Electron.* 2015;18:1–10https://doi.org/10.26879/506.
68. Maté-González M, Aramendi J, González-Aguilera D, Yravedra J. Statistical comparison between low-cost methods for 3D characterization of cut-marks on bones. *Remote Sens.* 2017;9:873https://doi:10.3390/rs9090873.
69. Edwards J, Rogers T. The accuracy and applicability of 3D modeling and printing blunt force cranial injuries. *J Forensic Sci.* 2017:1–9https://doi:10.1111/1556-4029.13627.
70. Viciano J, López-Lázaro S, Pérez-Fernández Á, Amores-Ampuero A, D'Anastasio R, Jiménez-Triguero JM. Scheuermann's disease in a juvenile male from the late Roman necropolis of Torrenueva (3rd-4th century CE, Granada, Spain). *Int J Paleopathol.* 2017;18:26–37https://doi.org/10.1016/j.ijpp.2017.04.003.
71. Porter ST, Roussel M, Soressi M. A simple photogrammetry rig for the reliable creation of 3D artifact models in the field: lithic examples from the early upper paleolithic sequence of les cottés (France). *Adv Archaeol Pract.* 2016;4:71–86https://doi.org/10.7183/2326-3768.4.1.71.
72. Massinon M, Dumont B, De Cock N, Salah SOT, Lebeau F. Study of retention variability on an early growth stage herbaceous plant using a 3D virtual spraying model. *Crop Prot.* 2015;78:63–71https://doi.org/10.1016/j.cropro.2015.08.018.
73. Secher JJ, Darvann TA, Pinholt EM. Accuracy and reproducibility of the DAVID SLS-2 scanner in three-dimensional facial imaging. *J Cranio-Maxillo-Fac Surg.* 2017;45:1662–1670https://doi.org/10.1016/j.jcms.2017.07.006.
74. Yravedra J, Aramendi J, Maté-González MÁ, Courtenay LA, González-Aguilera D. Differentiating percussion pits and carnivore tooth pits using 3D reconstructions and geometric morphometrics. *PLoS One.* 2018;13:e0194324https://doi.org/10.1371/journal.pone.0194324.
75. Zaimovic-Uzunovic N, Lemes S. Influences of surface parameters on laser 3D scanning. *IMEKO Conference Proceedings: International Symposium on Measurement and Quality Control.* 2010; 2010.
76. Perrone RV, Williams JL. Dimensional accuracy and repeatability of the NextEngine laser scanner for use in osteology and forensic anthropology. *J Archaeol Sci Rep.* 2019;25:308–319. https://doi.org/10.1016/j.jasrep.2019.04.012.
77. Veneziano A, Landi F, Profico A. Surface smoothing, decimation, and their effects on 3D biological specimens. *Am J Phys Anthropol.* 2018;166:473–480https://doi:10.1002/ajpa.23431.
78. Sholts SB, Wärmländer SKTS, Flores LM, Miller KWP, Walker PL. Variation in the measurement of cranial volume and surface area using 3d laser scanning technology. *J Forensic Sci.* 2010;55:871–876https://doi:10.1111/j.1556-4029.2010.01380.x.
79. Singh R, Baby B, Suri A, Anand S. Comparison of laser and structured light scanning techniques for neurosurgery applications. *3rd International Conference on Signal Processing and Integrated Networks (SPIN).* IEEE; 2016:301–305https://doi:10.1109/SPIN.2016.7566708.
80. Camison L, Bykowski M, Lee WW, et al. Validation of the Vectra H1 portable three-dimensional photogrammetry system for facial imaging. *Int J Oral Maxillofac Surg.* 2018;47:403–410https://doi:10.1016/j.ijom.2017.08.008.
81. Novotný V. Sex determination of the pelvic bone: a systems approach. *Anthropologie.* 1986;XXIV:197–206.
82. Gaillard J. Détermination sexuelle d'un os coxal fragmentaire. *Bull Mem Soc Anthropol Paris.* 1960;2:255–267.
83. Braüer G. Osteometrie. In: Knussmann R, ed. *Anthropol. Handb. Des Vergleichenden Biol. Des Menschen, Band 1.* Stuttgart: Gustav Fischer Verlag; 1988:160–232.
84. Schulter-Ellis F, Schmidt D, Hayek L, Craig J. Determination of sex with a discriminant analysis of new pelvic bone measurements: Part I. *J Forensic Sci.* 1983;28:169–180.

## Appendix A.4

International Journal of Legal Medicine
https://doi.org/10.1007/s00414-018-1934-1

ORIGINAL ARTICLE

# A validation study of the Stoyanova et al. method (2017) for age-at-death estimation quantifying the 3D pubic symphyseal surface of adult males of European populations

Anežka Kotěrová[1] · Jana Velemínská[1] · Eugénia Cunha[2] · Jaroslav Brůžek[1]

Received: 16 July 2018 / Accepted: 5 September 2018
© Springer-Verlag GmbH Germany, part of Springer Nature 2018

**Abstract**
The age-at-death estimation thresholds have recently been shifted towards a more objective assessment of the aging process. Such a non-subjective approach offers quantitative methods of age estimation; for instance, the method relating to the surfaces of pubic symphyses of males published by Stoyanova et al. (J Forensic Sci 62:1434–1444, 2017). A validation study was conducted to test the method performance in European samples. The sample consisted of 96 meshes of pubic symphyses of male individuals (known sex and age) that came from four different samples (two Portuguese collections, one Swiss, and one Crete). Stoyanova's method based on five regression models (three univariate and two multivariate models) performed worse in our sample, but only when the whole sample (without age limitation) was included. A sample limited to individuals under 40 years of age achieved better results in our study. The best results were reached through the thin plate spline algorithm (TPS/BE) with a root mean square error of 5.93 years and inaccuracy of 4.47 years. Generally, the multivariate regression models did not contribute to better age estimation. In our sample in all age categories, age was systematically underestimated. The quantitative method tested in this study works best for individuals under 40 years of age and provides a suitable basis for further research.

**Keywords** Pubic symphysis · Age-at-death · Quantitative analysis · 3D model · Structured light scanning

## Introduction

Age-at-death estimation from skeletal remains has contributed both in forensic investigation and bioarcheology, to the reconstruction of biological profiles in individual identification, and to the estimation of demographic profiles of past populations [1–3], respectively. Due to the differences between biological and chronological age, the age estimation of skeletonized adult remains is extremely problematic, especially in older individuals [4]. It is given by the nature of human senescence, which is based on the degenerative changes of articular surfaces, or dental wear [5–8].

Among all the skeletal age indicators, probably the greatest attention is generally paid to pubic symphysis. A large number of studies published between 1920 and 2018 serve as evidence (e.g., [9–18]). The most known and commonly used method of pubic symphysis for age-at-death estimation is the one proposed by Brooks and Suchey [12]. It is a phase-based method whereby the symphyseal surface is visually scored into six phases. Since the Suchey-Brooks method (often called the Suchey-Brooks system, SBS) was published in 1990, a considerable amount of revisions and tests have been done on it, both on dry bones (e.g., [15, 19–23]) and on virtual models [24–28]. Nowadays, it is known that degenerative changes on the pubic symphysis reflect increasing age only up to approximately 40 years of age (e.g., [14, 16, 20, 29–36]); in older individuals, the estimation becomes highly inaccurate. The main features used in age estimation are the appearance and

Jana Velemínská, Eugénia Cunha and Jaroslav Brůžek contributed equally to this work.

✉ Anežka Kotěrová
koterova.a@seznam.cz

[1] Department of Anthropology and Human Genetics, Faculty of Science, Charles University, Viničná 7, 128 43 Prague 2, Czech Republic

[2] Laboratory of Forensic Anthropology, Centre for Functional Ecology, Department of Life Sciences, University of Coimbra, Calçada Martim de Freitas, 3000-456 Coimbra, Portugal

Published online: 15 September 2018

fusion of the ventral rampart, where the fusion would be complete by 35 years of age [37]. Given that fact, some researchers use the pubic symphysis in their methods only up to 40 years [38]. For instance, the two-step procedure proposed by Baccino et al. [39, 40] combines two methods. The method of Suchey and Brooks [12] and the Lamendin dental method [41, 42] were used, where the first mentioned is used alone only up to the third phase (which approximately corresponds with an upper age of 40 years). Three other SBS phases (phase IV, V, and VI) are replaced by the Lamendin method, which evaluates age best between 40 and 60 years of age. Since no single indicator is able to capture the whole span of human life [43, 44], Xanthopoulou et al. [43] recommend the use of different age indicators for various age classes; for instance, pubic symphysis only for younger adults.

The known subjectivity of visual methods [4, 16, 45, 46] is another problem. A study of Kotěrová et al. (2018) confirms that despite the very sophisticated treatment of visually evaluated data (pubic symphysis and the auricular surface of the pelvic bone) in a multiethnic sample, these sophisticated approaches did not show any improvement in the accuracy of age-at-death estimation. The researchers assume that further designing or modifying existing methods, where the data are evaluated visually, will not fundamentally contribute to accurate and precise estimation [47]. Advanced technologies, such as 3D technologies (e.g., surface scanning, computed tomography scanners), have been recently set up in biological and forensic anthropology and are bringing new possibilities [48, 49].

The first attempt to quantify pubic symphyseal surfaces was published by Biwasaka et al. [28]. A curvature analysis was also later performed in a study of Villa et al. [50]. One of the most recent approaches is the computing of Dirichlet normal energy (DNE), as it was applied on pubic symphysis by Stock et al. [51]. DNE is a technique used for quantifying surfaces by using the changes in normal vectors in order to characterize general surface curvature. This topographic metric was initially used for tooth surfaces [52–54]. The advantage of DNE is that it is not affected by size (scale), position, or orientation [52, 54, 55]. Surfaces were analyzed in R package molaR [54]. The authors concluded that DNE expresses surface changes that could be used for age-at-death estimation; nevertheless, further research is required. Unfortunately, in regard to the above-mentioned approaches, no user-friendly software for age-at-death estimation has been proposed yet.

Another new approach for the fully quantitative evaluation of pubic symphysis surfaces based on the analysis of 3D surface scans was published by Stoyanova et al. [16, 17]. The analysis consists of a TPS/BE analysis (thin plate spline algorithm computing bending energy), a SAH-Score (*Slice and Algee-Hewitt*) algorithm [56]), and VC (ventral curvature) analysis. Unlike the above-mentioned approaches, Stoyanova proposes user-friendly software called "forAGE," which was first developed and presented in 2015 [16], with an extended version in 2017 [17]. Altogether, five regression models are computed, of which two are multivariate regression models (SAH-Score and VC, TPS/BE, and VC).

The purpose of the present study was to test and validate the quantitative shape analysis methods for age-at-death estimation proposed by Stoyanova et al. [17] on our sample of 3D models of pubic symphyses (age range 20–83 years). Additionally, we decided to test whether better results would be achieved in an age-limited sample. Our goal corresponds to the requirements of Daubert's criteria that the proposed methods be testable, tested, and accepted by the scientific community [57, 58]; especially in forensic practice, it is necessary to validate the methods. Due to the insufficient testing of these new methods, their use is limited [23].

## Material and methods

### Material

The dataset consists of 96 3D models of pelvic bones of adult male individuals from four European identified skeletal collections. Two of them are from Coimbra, Portugal, and are stored at the University of Coimbra. (1) The Coimbra Identified Skeletal Collection (CISC), assembled by Professor Tamagnini [59], represents individuals who were born between 1817 and 1924, and who died between 1904 and 1938. (2) The 21st Century Identified Skeletal Collection (CEI/XXI) [60] is a modern collection that consists of individuals who died between 1995 and 2009 (updated at the time of studying the collection). (3) The Heraklion Collection is a Cretan collection, which is housed at the facilities of the Forensic Pathology Division of the Hellenic Ministry of Justice and Human Rights in Crete, Greece. This collection consists of individuals born between 1867 and 1956, and died between 1968 and 1998 [61, 62]. (4) The last collection, the Simon Identified Skeletal Collection, is housed at the Laboratory of Prehistoric Archaeology and Anthropology of the University of Geneva, Switzerland [63, 64]. The Simon Collection includes individuals who lived in the second half of the twentieth century.

Only pelvic bones of individuals without obvious pathologies or taphonomy effects on the pubic symphysis surface were selected, and only those of male individuals (as in the original study of Stoyanova et al. [16, 17]). The left side was used preferentially; only if it was absent or damaged was the right side used. Detailed descriptions of the samples are presented in Tables 1 and 2. The age of our samples ranged from 20 to 83 years, with a mean age of 45.13 years. The whole sample was further divided into two subgroups; the first consisted of individuals under 40 years ($n = 41$), and the second was comprised of individuals over 40 years ($n = 55$). The

**Table 1** Summary table of osteological collections used in the present study; the numbers of male individuals in six age categories

| | Number of individuals in each age category | | | | | | |
|---|---|---|---|---|---|---|---|
| Collection | ≤20 | 21–30 | 31–40 | 41–50 | 51–60 | 61< | Total |
| Coimbra 1 | 0 | 14 | 6 | 19 | 10 | 11 | 60 |
| Coimbra 2 | 0 | 3 | 4 | 2 | 1 | 1 | 11 |
| Geneva | 1 | 3 | 7 | 1 | 2 | 2 | 16 |
| Heraklion | 0 | 0 | 3 | 0 | 3 | 3 | 9 |
| | | | | | | | 96 |
| | | | | | | Mean age (years) = | 45.13 |

limit of 40 years was chosen since it is in related literature considered the maximal age where the surface changes of the pubic symphysis continue to correlate with age [14, 20, 29, 30, 32, 33].

## Methods

### Data acquisition, manipulation, and analysis

The tested method of Stoyanova et al. (2017) is based on surface meshes; thus, 3D models of pelvic bones were made as a first step. All the 3D models in the present study were made by the David SLS-2 surface scanner (later renamed to HP 3D Structured Light Scanner PRO 2) and saved in ply format. Whole pelvic bones were scanned and subsequently post-processed (multiple scans were aligned to form a single mesh). Integrated software (David LaserScanner v.3.10.4) was used to take individual scans, for post-processing, and then to create the final 3D models. The isolation of the articulation surfaces of the pubic symphyses from the rest of the bone (Fig. 1) was performed using MeshLab software [65]. Isolated pubic symphysis mesh in ply format was analyzed with forAge software (http://morphlab.sc.fsu.edu/software/forAge/). Each mesh had to be simplified (we performed consistent simplification of 15,000 faces) before loading into the software. Firstly, the loaded scan was translated, scaled, and rotated with the use of principal component analysis (PCA), which was done automatically. After that, the users had to manually flip the scan so the ventral margin of the pubic symphysis was facing upward (viz., forAge User Manual).

Age-at-death estimates were produced using five regression models: the TPS/BE analysis, SAH-Score (Slice and Algee-Hewitt), VC (ventral curvature), and a combination of VC with another two analyses. The first two analytical methods capture the flatness of the surface: the thin plate spline algorithm (TPS) measures the bending energy (BE), and the SAH-Score is an algorithm published by Slice and Algee-Hewitt [56]. The last method is shape measurement based on algorithms that capture age-related changes of ventral margin curvature.

**Table 2** Numbers of male individuals in two age groups: under 40 and over 40 with mean ages

| | Number of individuals in each age category | |
|---|---|---|
| Collection | ≤40 (years) | 40< (years) |
| Coimbra 1 | 20 | 40 |
| Coimbra 2 | 7 | 4 |
| Geneva | 11 | 5 |
| Heraklion | 3 | 6 |
| Total | 41 | 55 |
| Mean age (years) | 30.41 | 56.90 |

### Statistical analyses

All results are provided for all the data together, as well as separately for individuals up to 40 years and individuals over 40 years.

Three characteristics were used to compare the regression models with each other: firstly, bias, which expresses the mean overestimation and underestimation of the individual's actual age; secondly, inaccuracy, which measures the average absolute error of age estimation; and, finally, root mean square error, which is the square root of the average of the squared differences between the estimated and actual values.

The bias and inaccuracy were calculated using the following formulae, as they were used in the original study of Stoyanova et al. [17] and other studies as well (e.g., [29, 43]):

$$\text{Bias} = \sum(\text{Estimated age} - \text{Actual age})/N$$

$$\text{Inaccuracy} = \sum|\text{Estimated age} - \text{Actual age}|/N$$

where $N$ is the sample size.

The root mean square error (RMSE) is defined as

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(y_j - \hat{y}_j\right)^2},$$

where $n$ is the sample size, $y_j$ is the estimated age, and $\hat{y}_j$ is the actual age.

A paired $t$ test was applied to access the significance of the relationship between the actual and estimated ages for all five models in three samples (the whole dataset, the under 40 years, and the over 40 years). To avoid the problem of multiple comparisons, the Bonferroni correction to the $p$ value was done. A two-sample $t$ test was applied to evaluate how the method estimates in the samples up to 40 years and over 40 years. Statistical analyses were performed using the

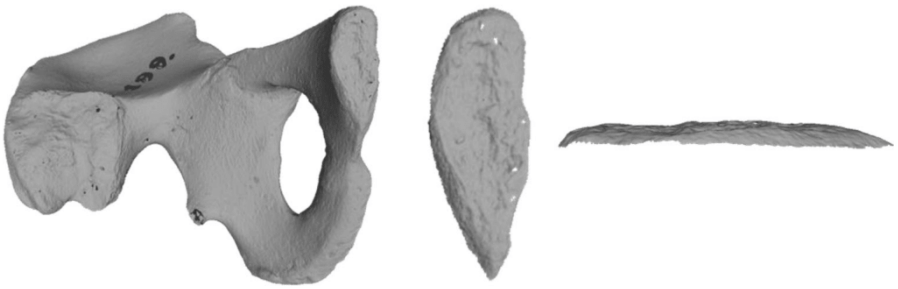

**Fig. 1** An example of a 3D scan of the whole pelvic bone (left) and an isolated pubic symphyseal surface (in the middle and right) as a smooth surface

statistical program "R" (http://www.r-project.org) and in Microsoft Excel 2007.

## Results

The bias and inaccuracy as well as RMSE and the results of paired *t* tests of all five regression models were computed for the whole sample, for individuals ≤40 years and individuals >40 years, and are shown in Table 3.

The bias and inaccuracy for the whole dataset ranged between −15.67 (TPS/BE + VC) and −8.43 (VC) years, and 14.15 (VC) and 16.96 (TPS/BE + VC) years, respectively. The RMSE values were from 18.35 (VC) to 22.25 (TPS/BE + VC) years. In the case of the sample containing only individuals ≤40 years, the results of bias and inaccuracy were between −1.77 (TPS/BE + VC) and 5.90 years (VC) years, and 4.47 (TPS/BE) and 7.07 (VC) years, respectively. The values of RMSE reached 5.93 years at the minimum (TPS/BE) and 9.49 years at the maximum (VC). The bias and inaccuracy for the sample limited to individuals over 40 years of age were from −26.04 (TPS/BE + VC) to −19.11 (VC) years, and 19.42 (VC) to 26.04 (TPS/BE + VC) years, respectively. The RMSE ranged from 22.82 (VC) to 28.88 (TPS/BE + VC) years. The differences between the actual and estimated ages were statistically significant ($p$ value < 0.05 after Bonferroni correction) for all models in the whole dataset and in the sample of individuals older than 40 years. In the samples limited to 40 years, the differences were statistically significant only in the case of the VC ($p$ value < 0.001). For the remaining four regression models, TPS/BE ($p$ value = 0.38), SAH ($p$ value = 1), TPS/BE + VC ($p$ value = 0.37), and SAH + VC ($p$ value = 1), there were no significant differences between the actual and estimated ages (at an alpha level of 0.05). All $p$ values are corrected using Bonferroni correction.

**Table 3** Results of bias, inaccuracy, RMSE, and paired *t* test for all five tested models. Results for the whole sample and samples under and over 40 years

| Regression model | Age category | RMSE (years) | Bias | Inaccuracy | *p* value | |
|---|---|---|---|---|---|---|
| | | | | | Uncorrected | Corrected |
| TPS/BE | All data | 22.09 | −15.54 | 16.75 | < 0.001 | < 0.001 |
| | ≤40 years | 5.93 | −1.64 | 4.47 | 0.08 | 0.38* |
| | >40 years | 28.74 | −25.91 | 25.91 | < 0.001 | < 0.001 |
| SAH | All data | 20.91 | −13.58 | 15.70 | < 0.001 | < 0.001 |
| | ≤40 years | 7.34 | −0.78 | 5.75 | 0.50 | 1* |
| | >40 years | 26.88 | −23.13 | 23.13 | < 0.001 | < 0.001 |
| VC | All data | 18.35 | −8.43 | 14.15 | < 0.001 | < 0.001 |
| | ≤40 years | 9.49 | 5.90 | 7.07 | < 0.001 | < 0.001 |
| | >40 years | 22.82 | −19.11 | 19.42 | < 0.001 | < 0.001 |
| TPS/BE + VC | All data | 22.25 | −15.67 | 16.96 | < 0.001 | < 0.001 |
| | ≤40 years | 6.37 | −1.77 | 4.79 | 0.07 | 0.37* |
| | >40 years | 28.88 | −26.04 | 26.04 | < 0.001 | < 0.001 |
| SAH + VC | All data | 21.08 | −13.67 | 15.99 | < 0.001 | < 0.001 |
| | ≤40 years | 7.48 | −0.70 | 6.13 | 0.56 | 1* |
| | >40 years | 27.09 | −23.34 | 23.34 | < 0.001 | < 0.001 |

*p* values are shown for a two-tailed test. The last column gives the *p* value after the Bonferroni correction for multiple comparisons. Non-significant differences between actual and estimated ages are marked with asterisks

*TPS/BE* thin plate spline/bending energy, *SAH* Slice and Algee-Hewitt score, *VC* ventral curvature

The results of a two-sample *t* test (Table 4), which was applied to reveal the differences between the actual and estimated age values, show that all five regression models estimate the age in the two age subgroups (under 40 years and over 40 years) significantly differently (*p* values < 0.05).

A summary of the results of RMSE, bias, and inaccuracy values (for the whole dataset and sample limited up to 40 years) from the original study of Stoyanova et al. (2017) and from the present study for each method is provided in Table 5. RMSE values for the age category 20–40 years (adopted by Stoyanova from Miranker [66] for better comparison of results) are not provided.

## Discussion

This is the first study where the quantitative aging method of Stoyanova et al. [17] has been validated and tested using European populations. Four European skeletal collections were used: the Coimbra Identified Skeletal Collection, the twenty-first Century Identified Skeletal Collection, the Cretan Collection, and the Simon Collection. Altogether, we analyzed 96 meshes of pubic symphyses of male individuals. In addition to validating and testing the methods across the whole sample, we also wanted to compare how the methods would estimate age in the samples up to 40 years ($n = 41$) and over 40 years ($n = 55$). To make these two samples comparable, approximately the same numbers in subsamples were required. The total number of individuals was limited for several reasons. First of all, there are generally small numbers of younger individuals (< 40 years) in the collections. For instance, there were only 10 individuals under 40 years, of which 8 were male (total $n = 247$, update in the time of studying the collection). The collections have a higher number of individuals who died at an older age. Secondly, there was considerable damage or even the complete absence of pubic symphyses caused by taphonomic influences [8], which is very common. However, our sample size is comparable to the sample used by Stoyanova et al. [17].

**Table 4** Results of the two-sample *t* tests between two subgroups (samples under and over 40 years)

| | *p* value | |
|---|---|---|
| Regression model | Uncorrected | Corrected |
| TPS/BE | < 0.001 | < 0.001 |
| SAH | < 0.001 | < 0.001 |
| VC | < 0.001 | < 0.001 |
| TPS/BE + VC | < 0.001 | < 0.001 |
| SAH + VC | < 0.001 | < 0.001 |

*p* values of a two-sample *t* test before and after Bonferroni correction. All *p* values are highly statistically significant

*TPS/BE* thin plate spline/bending energy, *SAH* Slice and Algee-Hewitt score, *VC* ventral curvature

In the case of age-at-death estimation, the new approach on the rise is the mathematical quantification of the articular surfaces of pelvic articulation areas. Different approaches are being applied in order to quantitatively evaluate the articular surface changes that help in estimating age. At the forefront are the age indicators on the os coxae, mainly the pubic symphysis [16, 17, 28, 50, 51], and the auricular surface of the ilium [50]. Some publications that deal with the evaluation of age-at-death in digital data are summarized in the review article by Villa et al. in 2016 [49].

Software called forAGE was first presented in 2015 [16]. The analysis of three-dimensional surface scans of the pubic symphysis ($n = 44$ male individuals and 12 casts) in this first version included only a thin plate spline algorithm (TPS) that measures bending energy (BE). In the same year, Stoyanova's colleagues Slice and Algee-Hewitt [56] presented a paper where the SAH-Score (*Slice and Algee-Hewitt*) was introduced. Two years later, in 2017, this team [17] published a paper whereby the TPS analysis, SAH-Score, and, moreover, the shape measurements that capture age-related changes of ventral margin curvature (VC) were combined and incorporated into the forAGE software. Their dataset consisted of 68 male individuals and 25 casts (casts for the implementation of the Suchey-Brooks system and McKern and Stewart's component-based scoring system). The age-at-death range of the sample was 16–90 years, with an emphasis on the equal distribution of individuals in pre-middle-aged and post-middle-aged subgroups, where the sectioning point was 40 years. Results from this last study significantly improved those from the previous studies of this particular author and team [16, 56]. The selection of only male individuals was explained by the authors through the documented variation in pubic symphyses between the sexes, and even among different populations. On the contrary, there were several studies published that did not find sex differences in the aging patterns (e.g., [34, 67]). However, it is obvious that in the future such regression models should be designed for females as well.

The results of our study clearly show, and confirm plenty of various studies on pubic symphysis (e.g., [32, 33, 36]), that pubic symphysis as an age indicator cannot be used for the whole of a lifespan. The results of bias, inaccuracy, and RMSE speak for themselves. When using all the data (without age limitation), the lowest values of RMSE, bias, and inaccuracy were reached by VC (18.35, − 8.43, and 14.15 years, respectively). The highest values, on the contrary, showed TPS/BE + VC (22.25, − 15.67, and 16.96 years, respectively). Even higher values were achieved when samples of individuals over 40 years of age were used, again with the lowest values for VC (RMSE 22.82, bias − 19.11, and inaccuracy 19.42 years) and highest for TPS/BE + VC (28.88, − 26.04,

**Table 5** A summary table comparing the values (in years) of RMSE, bias, and inaccuracy for each method. Results for the whole dataset and the sample limited to 20–40 years in the original study of Stoyanova et al. (2017) and in the present study are provided

| Regression model | | Whole dataset | | 20–40 years | |
|---|---|---|---|---|---|
| | | Kotěrová et al. | Stoyanova et al. [17] | Kotěrová et al. | Stoyanova et al. [17] |
| TPS/BE | RMSE | 22.09 | 16.38 | 5.93 | – |
| | Bias | − 15.54 | − 2.51 | − 1.64 | 8.47 |
| | Inaccuracy | 16.75 | 12.58 | 4.47 | 9.08 |
| SAH | RMSE | 20.91 | 14.15 | 7.34 | – |
| | Bias | − 13.58 | − 1.96 | − 0.78 | 5.71 |
| | Inaccuracy | 15.7 | 10.81 | 5.75 | 8.1 |
| VC | RMSE | 18.35 | 16.55 | 9.49 | – |
| | Bias | − 8.43 | − 2.73 | 5.9 | 5.84 |
| | Inaccuracy | 14.15 | 12.86 | 7.07 | 9.89 |
| TPS/BE + VC | RMSE | 22.25 | 15.07 | 6.37 | – |
| | Bias | − 15.67 | − 2.21 | − 1.77 | 7.76 |
| | Inaccuracy | 16.96 | 11.39 | 4.79 | 8.93 |
| SAH + VC | RMSE | 21.08 | 13.68 | 7.48 | – |
| | Bias | − 13.67 | − 1.82 | − 0.7 | 5.76 |
| | Inaccuracy | 15.99 | 10.79 | 6.13 | 8.36 |

RMSE for age category 20–40 years in Stoyanova et al. [17] was not calculated

*TPS/BE* thin plate spline/bending energy, *SAH* Slice and Algee-Hewitt score, *VC* ventral curvature

and 26.04 years, respectively). Individual values often differed only slightly, but in all cases, there were statistically significant differences between the estimated and actual ages. As expected, the best results were reached with samples limited to the age of 40. Here, the highest values of all three characteristics were shown by VC with RMSE equal to 9.49 years, bias to 5.90 years, and inaccuracy to 7.07 years. The lowest values, on the other hand, were 5.93 years (TPS/BE) for RMSE; in the case of bias, it was − 0.70 years (SAH + VC); and for inaccuracy, it was 4.47 years (TPS/BE). Apart from the regression model, which evaluates ventral curvature (VC), the differences between actual and estimated ages were not significant.

Based on the overall results of RMSE, bias, and inaccuracy in the whole sample, Stoyanova recorded lower values for all three characteristics than we have shown here; thus, they presented slightly better estimates. RMSE ranged between 13.68 and 16.55 years; bias was between − 1.82 and − 2.73 years; and, finally, values of inaccuracy ranged between 10.79 and 12.86 years. Given that fact, it was more surprising that this is not true for the sample of up to 40 years as well. Here, on the contrary, lower values were found in our study. Although we could not compare RMSE values for the age category of 20–40 years (since they are not provided in the study of Stoyanova), our bias and inaccuracy results were lower by 4.93–6.83 and 2.23–4.61 years, respectively, with an exception of bias for VC, which is almost the same in both studies. Negative values of bias were observed in our data in the whole dataset, as well as in the samples under 40 years and over 40. This indicates an underestimation of the actual age, while the only positive value of bias in the case of VC in the sample under 40 years shows a more likely overestimation of age. Stoyanova consistently overestimated the individuals under 40 years and underestimated those over 60 years. Stoyanova et al. compare their results against the performance of the Suchey-Brooks method, and concluded that their models produced generally smaller age intervals (2× RMSE, ranged from about 27 to 32 years) [17]. As reported above, our age intervals when the whole human life span is considered are broader (since the RMSE values are higher). For individuals up to 40 years, the age intervals in our study were more promising (from ± 6 to 7.5 years, about 12 to 15 years, respectively) when the VC value was omitted. The question about the suitability of RMSE as a characteristics for model evaluation has to be raised, since the root mean square error is sensitive to outliers [68] in contrast to, for example, the level of inaccuracy.

The study of Stoyanova et al. reported that the curvature of the ventral margin itself did not outperform any of the surface measures. However, in their sample, VC contributed to better estimation when combined with both TPS/BE and SAH in multivariate regression models. This, however, is not the case for our specifically European sample. We observed even slightly higher values (in decimal numbers) of all the assessed characteristics (RMSE, bias, and inaccuracy) when TPS/BE or SAH was combined with VC. This is true for all age

categories (i.e., the whole dataset and samples under and over 40 years). The only small exceptions are the bias of SAH and SAH + VC (− 0.78 and − 0.70 years, respectively). One of the possible causes may be the fact that the symphyseal articular surface must be selected manually from the rest of the mesh by a researcher; therefore, there may be differences between the researchers in this very essential step. Unfortunately, the symphyseal face is not always well delimited; in some cases, the use of texture could help in making a decision, but sometimes it depends only on the researcher. For that reason, an inter-observer error should be evaluated in the future. Another possible cause for the differences in the contribution of VC to a more accurate estimation between the tested study of Stoyanova and our study could be the various resolutions of scans. Stoyanova states that the manual selection mentioned before can also partially depend on the resolution.

A NextEngine 3D Desktop laser scanner (http://www.nextengine.com/) was used to digitize the sample data in the study of Stoyanova et al. Laser scanners [16, 17, 50] or CT scanners [24, 50, 51] are the most commonly used devices to capture 3D surfaces of regions of interest. However, in the present study, we used a less known structured light scanner (HP 3D Structured Light Scanner Pro S2), which uses projected light patterns and a camera system to capture the three-dimensional shape of an object. This particular scanner (previously known as David SLS) has already been used in several studies for scanning both human and animal bones [69–72], as well as for other purposes [73–75]. Recently, the accuracy and reproducibility of the scanner were tested on mannequin faces, as well as those of live subjects [76]. Some undeniable advantages of this scanner are the low scanning time, portability, high resolution, and low cost. Although it is not yet known whether the outputs from both of these scanners are comparable, the preliminary results (not published yet) of our study suggest that the results of the HP 3D Structured Light Scanner Pro S2 are even more accurate.

The question of population specificity must also be discussed here. The original study is based on solely American samples of white males, while we used males from several European populations. Although the population specificity of the aging methods was demonstrated several times (e.g., [19, 77–79]), there are some studies that did not confirm it; for instance, Kimmerle et al. [80] did not find population differences for males, nor did Sakaue [20]. At the same time, it must be stressed here that globalization and migration irretrievably change the composition of populations. If there are no indicators as to which population a set of particular human remains belongs, we cannot assign them with certainty. These remains do not necessarily originate from the population of the place of finding [40, 81–83]. Methods derived from multipopulation samples could be a solution [84].

Quantitative methods have a potential to outperform the traditional visual methods, since they are able to reduce the subjectivity and could possibly reveal an aging pattern that is not possible to determine through visual observation. Nevertheless, in the case of pubic symphysis, an upper limit of approximately 40 years remains the maximal value for accurate age estimation. Future efforts should be concentrated on using the various age indicators that emphasize their biological ability to reflect the aging process.

## Conclusion

This study aimed to test a recent quantitative aging method on pubic symphyses for male individuals from European samples. The method of Stoyanova et al. [17] performed better in our sample (lower values of RMSE, bias, and inaccuracy were reached) when individuals under 40 years were analyzed. Apart from VC, the values of RMSE ranged between 5.93 and 7.48 years. In a sample comprising individuals without age limitations, the previous results of many studies were confirmed: the use of pubic symphysis for individuals older than 40 years leads to very inaccurate estimates with overly broad age intervals (± 18.35–± 22.35 years). A computational method that measures the curvature of the ventral margin did not improve age estimation when combined into multivariate regression models.

In conclusion, a quantitative evaluation of surface changes related to age provides an objective assessment. However, the biological possibilities of individual skeletal indicators to reflect age changes should be respected in further attempts to quantify surface changes.

**Acknowledgements** We would like to thank the Department of Life Sciences, University of Coimbra, for providing access to the 21st Century Identified Skeletal Collection and the Coimbra Identified Skeletal Collection (CISC). We also thank Dr. Elena Kranioti for providing access to the Cretan Collection and Ms. Lada Šťovíčková for digitizing it. Last but not least, we would like to thank Dr. Jocelyn Desideri and those at the Laboratory of Prehistoric Archaeology and Anthropology of the University of Geneva for providing access to the Simon Collection, and to Dr. François Marchal for providing the David SLS-2 surface scanner. Finally, we would like to thank to Ms. Šárka Roušavá for the linguistic corrections.

**Funding information** This research has been supported by the research grant GAUK No. 642218.

## References

1. Hoppa RD, Vaupel JW (2002) The Rostock Manifesto for paleodemography: the way from stage to age. In: Hoppa RD, Vaupel JW (eds) Paleodemography: age distributions from skeletal samples. Cambridge University Press, Cambridge, pp 1–8

2. Baccino E, Schmitt A (2006) Determination of adult age at death in the forensic context. In: Schmitt A, Cunha E, Pinheiro J (eds) Forensic anthropology and medicine: complementary sciences from recovery to cause of death. Humana Press Inc., Totowa, pp 259–280
3. Cattaneo C (2007) Forensic anthropology: developments of a classical discipline in the new millennium. Forensic Sci Int 165:185–193
4. Kemkes-Grottenthaler A (2002) Aging through the ages: historical perspectives on age indicator methods. In: Hoppa R, Vaupel J (eds) Paleodemography: age distributions from skeletal samples. Cambridge University Press, Cambridge, pp 48–72
5. Overbury RS, Cabo LL, Dirkmaat DC, Symes SA (2009) Asymmetry of the os pubis: implications for the Suchey-Brooks method. Am J Phys Anthropol 139:261–268
6. Passalacqua NV (2009) Forensic age-at-death estimation from the human sacrum. J Forensic Sci 54:255–262
7. Buckberry J (2015) The (mis)use of adult age estimates in osteology. Ann Hum Biol 42:323–331
8. Cappella A, Cummaudo M, Arrigoni E, Collini F, Cattaneo C (2017) The issue of age estimation in a modern skeletal population: are even the more modern current aging methods satisfactory for the elderly? J Forensic Sci 62:12–17
9. Todd WT (1920) Age changes in the pubic bone. I. The male white pubis. Am J Phys Anthropol 3:285–334
10. Gilbert BM, McKern TW (1973) A method for aging the female os pubis. Am J Phys Anthropol 38:31–38
11. Katz D, Suchey JM (1989) Race differences in pubic symphyseal aging patterns in the male. Am J Phys Anthropol 80:167–172
12. Brooks S, Suchey JM (1990) Skeletal age determination based on the os pubis: a comparison of the Acsádi- Nemeskéri and Suchey-Brooks methods. Hum Evol 5:227–238
13. Chen X, Zhang Z, Tao L (2008) Determination of male age at death in Chinese Han population: using quantitative variables statistical analysis from pubic bones. Forensic Sci Int 175:36–43
14. Hanihara K, Suzuki T (1978) Estimation of age from the pubic symphysis by means of multiple regression analysis. Am J Phys Anthropol 48:233–240
15. Hartnett KM (2010) Analysis of age-at-death estimation using data from a new, modern autopsy sample—part I: pubic bone. J Forensic Sci 55:1145–1151
16. Stoyanova D, Algee-Hewitt BFB, Slice DE (2015) An enhanced computational method for age-at-death estimation based on the pubic symphysis using 3D laser scans and thin plate splines. Am J Phys Anthropol 158:431–440
17. Stoyanova DK, Algee-Hewitt BFB, Kim J, Slice DE (2017) A computational framework for age-at-death estimation from the skeleton: surface and outline analysis of 3D laser scans of the adult pubic symphysis. J Forensic Sci 62:1434–1444
18. Savall F, Hérin F, Peyron PA, Rougé D, Baccino E, Saint-Martin P, Telmon N (2018) Age estimation at death using pubic bone analysis of a virtual reference sample. Int J Legal Med 132:609–615
19. Schmitt A (2004) Age-at-death assessment using the os pubis and the auricular surface of the ilium: a test on an identified Asian sample. Int J Osteoarchaeol 14:1–6
20. Sakaue K (2006) Application of the Suchey–Brooks system of pubic age estimation to recent Japanese skeletal material. Anthropol Sci 114:59–64
21. San Millán M, Rissech C, Turbón D (2013) A test of Suchey-Brooks (pubic symphysis) and Buckberry-Chamberlain (auricular surface) methods on an identified Spanish sample: paleodemographic implications. J Archaeol Sci 40:1743–1751
22. Berg GE (2008) Pubic bone age estimation in adult women. J Forensic Sci 53:569–577
23. Rivera-Sandoval J, Monsalve T, Cattaneo C (2018) A test of four innominate bone age assessment methods in a modern skeletal collection from Medellin, Colombia. Forensic Sci Int 282:232–2e1
24. Wink AE (2014) Pubic symphyseal age estimation from three-dimensional reconstructions of pelvic ct scans of live individuals. J Forensic Sci 59:696–702
25. Telmon N, Gaston A, Chemla P, Blanc A, Joffre F, Rougé D (2005) Application of the Suchey-Brooks method to three - dimensional imaging of the pubic symphysis. J Forensic Sci 50:1–6
26. Lottering N, MacGregor DM, Meredith M, Alston CL, Gregory LS (2013) Evaluation of the Suchey-Brooks method of age estimation in an Australian subpopulation using computed tomography of the pubic symphyseal surface. Am J Phys Anthropol 150:386–399
27. Tocheri MW, Razdan A, Dupras TL, Bae M, Liu D (2002) Three dimensional quantitative analyses of human pubic symphyseal morphology: can current limitations of skeletal aging methods be resolved? Am J Phys Anthropol (Suppl 34) 155
28. Biwasaka H, Sato K, Aoki Y, Kato H, Maeno Y, Tanijiri T, Fujita S, Dewa K (2013) Three dimensional surface analyses of pubic symphyseal faces of contemporary Japanese reconstructed with 3D digitized scanner. Legal Med 15:264–268
29. Meindl RS, Lovejoy CO, Mensforth RP, Walker RA (1985) A revised method of age determination using the os pubis, with a review and tests of accuracy of other current methods of pubic symphyseal aging. Am J Phys Anthropol 68:29–45
30. Lovejoy CO, Meindl RS, Tague R, Latimer B (1995) The senescent biology of the hominoid pelvis: its bearing on the pubic symphysis and auricular surface as age-at-death indicators in the human skeleton. Riv di Antropol 73:31–49
31. Lovejoy CO, Meindl RS, Tague R, Latimer B (1997) The comparative senescent biology of the hominoid pelvis and its implications for the use of age-at-death indicators in the human skeleton. In: Paine R (ed) Integrating archaeological demography: multidisciplinary approaches to prehistoric population. Southern Illinois University Press, Carbondale, pp 43–63
32. Schmitt A, Murail P, Cunha E, Rougé D (2002) Variability of the pattern of aging on the human skeleton: evidence from bone indicators and implication on age at death estimation. J Forensic Sci 47: 1203–1209
33. Márquez-Grant N (2015) An overview of age estimation in forensic anthropology: perspectives and practical considerations. Ann Hum Biol 42:308–322
34. Buk Z, Kordik P, Bruzek J, Schmitt A, Snorek M (2012) The age at death assessment in a multi-ethnic sample of pelvic bones using nature-inspired data mining methods. Forensic Sci Int 220:294 e1-e9
35. Rissech C, Estabrook GF, Cunha E, Malgosa A (2006) Using the acetabulum to estimate age at death of adult males. J Forensic Sci 51:213–229
36. Gocha TP, Ingvoldstad ME, Kolatorowicz A, Cosgriff-Hernandez MTJ, Sciulli PW (2015) Testing the applicability of six macroscopic skeletal aging techniques on a modern Southeast Asian sample. Forensic Sci Int 249:318–3e1
37. Meindl RS, Russell KF (1998) Recent advances in method and theory in paleodemography. Annu Rev Anthr 27:375–399
38. Dudzik B, Langley NR (2015) Estimating age from the pubic symphysis: a new component-based system. Forensic Sci Int 257:98–105
39. Baccino E, Zerilli A (1997) The two step strategy (TSS) or the right way to combine a dental (Lamendin) and an anthropological (Suchey–Brooks system) method for age determination. In: Proceedings of the American Academy of Forensic Sciences XLIX, p 150
40. Baccino E, Sinfield L, Colomb S, Baum TP, Martrille L (2014) Technical note: the two step procedure (TSP) for the determination

of age at death of adult human remains in forensic cases. Forensic Sci Int 244:247–251
41. Lamendin H (1988) Appréciation d'age par la méthode de Gustafson <<simplifiée>>. Chir Dent Fr 427:205–214
42. Lamendin H, Baccino E, Humbert JF, Tavernier JC, Nossintchouk RM, Zerilli A (1992) A simple technique for age estimation in adult corpses: the two criteria dental method. J Forensic Sci 37:1373–1379
43. Xanthopoulou P, Valakos E, Youlatos D, Nikita E (2018) Assessing the accuracy of cranial and pelvic ageing methods on human skeletal remains from a modern Greek assemblage. Forensic Sci Int 286:266–2e1
44. Martrille L, Ubelaker DH, Cattaneo C, Seguret F, Tremblay M, Baccino E (2007) Comparison of four skeletal methods for the estimation of age at death on white and black adults. J Forensic Sci 52:302–307
45. Galera V, Ubelaker D, Hayek L (1995) Interobserver error in macroscopic methods of estimating age at death from the human skeleton. Int J Anthr 10:229–239
46. Kimmerle EH, Prince DA, Berg GE (2008) Inter-observer variation in methodologies involving the pubic symphysis, sternal ribs, and teeth. J Forensic Sci 53:594–600
47. Kotěrová A, Navega D, Štepanovský M, Buk Z, Brůžek J, Cunha E (2018) Age estimation of adult human remains from hip bones using advanced methods. Forensic Sci Int 287:163–175
48. Algee-Hewitt BFB (2013) Age estimation in modern forensic anthropology. In: Tersigni-Tarrant MA, Shirley NR (eds) Forensic anthropology: an introduction. CRC Press, Boca Raton, pp 181–230
49. Villa C, Buckberry J, Lynnerup N (2016) Evaluating osteological ageing from digital data. J Anat. https://doi.org/10.1111/joa.12544
50. Villa C, Buckberry J, Cattaneo C, Frohlich B, Lynnerup N (2015) Quantitative analysis of the morphological changes of the pubic symphyseal face and the auricular surface and implications for age at death estimation. J Forensic Sci 60:556–565
51. Stock M, Morse P, Villa C (2017) Quantitative assessment of age-related topographic changes in the pubic symphysis. In: 86th Annual Meeting of the American Association of Physical Anthropologists, p 371
52. Bunn JM, Boyer DM, Lipman Y, St. Clair EM, Jernvall J, Daubechies I (2011) Comparing Dirichlet normal surface energy of tooth crowns, a new technique of molar shape quantification for dietary inference, with previous methods in isolation and in combination. Am J Phys Anthropol 145:247–261
53. Godfrey LR, Winchester JM, King SJ, Boyer DM, Jernvall J (2012) Dental topography indicates ecological contraction of lemur communities. Am J Phys Anthropol 148:215–227
54. Pampush JD, Winchester JM, Morse PE, Vining AQ, Boyer DM, Kay RF (2016) Introducing molaR: a new R package for quantitative topographic analysis of teeth (and other topographic surfaces). J Mamm Evol 23:397–412
55. Spradley JP, Pampush JD, Morse PE, Kay RF (2017) Smooth operator: the effects of different 3D mesh retriangulation protocols on the computation of Dirichlet normal energy. Am J Phys Anthropol 163:94–109
56. Slice DE, Algee-Hewitt BFB (2015) Modeling bone surface morphology: a fully quantitative method for age-at-death estimation using the pubic symphysis. J Forensic Sci 60:835–843
57. Daubert V. Merrell Dow Pharmaceuticals (1993), Inc., 509 US 579
58. Grivas CR, Komar DA (2008) Kumho, Daubert, and the nature of scientific inquiry: implications for forensic anthropology. J Forensic Sci 53:771–776
59. Cunha E, Wasterlain S (2007) The Coimbra identified osteological collections. In: Grupe G, Peters J (eds) Skeletal series in their socioeconomic context. Documenta Archaeobiologiae. Verlag Marie Leidorf, Rahden, pp 23–33
60. Ferreira MT, Vicente R, Navega D, Gonçalves D, Curate F, Cunha E (2014) A new forensic collection housed at the University of Coimbra, Portugal: the 21st century identified skeletal collection. Forensic Sci Int 245:202–2e1
61. Kranioti EF, Işcan MY, Michalodimitrakis M (2008) Craniometric analysis of the modern Cretan population. Forensic Sci Int 180:1–5
62. Kranioti EF, Michalodimitrakis M (2009) Sexual dimorphism of the humerus in contemporary cretans—a population-specific study and a review of the literature. J Forensic Sci 54:996–1000
63. Perreard Lopreno G (2007) Adaptation structurelle des os du membre supérieur et de la clavicule à l'activité. Doctoral dissertation, University of Geneva
64. Abegg C, Desideri J (2017) A probable case of multiple myeloma in a female individual from the Simon Identified Skeletal Collection (late 19th–early 20th century, Vaud, Switzerland). Int J Paleopathol 21:158–165. https://doi.org/10.1016/j.ijpp.2017.02.001
65. Cignoni P, Callieri M, Corsini M, Dellepiane M, Ganovelli F, Ranzuglia G (2008) MeshLab: an open-source mesh processing tool. In 6th Eurographics Ital Chapter Conf, pp 129–136
66. Miranker M (2016) A comparison of different age estimation methods of the adult pelvis. J Forensic Legal Med 61:1173–1179
67. Calce SE (2012) A new method to estimate adult age-at-death using the acetabulum. Am J Phys Anthropol 148:11–23
68. Chai T, Draxler RR (2014) Root mean square error (RMSE) or mean absolute error (MAE)?—Arguments against avoiding RMSE in the literature. Geosci Model Dev 7:1247–1250
69. Tambusso PS, Mcdonald HG, Fariña RA (2015) Description of the stylohyal bone of a giant sloth (Lestodon armatus). Palaeontol Electron 18:1–10
70. González MÁM, Yravedra J, González-Aguilera D, Palomeque-González JF, Domínguez-Rodrigo M (2015) Micro-photogrammetric characterization of cut marks on bones. J Archaeol Sci 62:128–142
71. Edwards J, Rogers T (2017) The accuracy and applicability of 3D modeling and printing blunt force cranial injuries. J Forensic Sci 63:683–691. https://doi.org/10.1111/1556-4029.13627
72. Viciano J, López-Lázaro S, Pérez-Fernández Á, Amores-Ampuero A, D'Anastasio R, Jiménez-Triguero JM (2017) Scheuermann's disease in a juvenile male from the late Roman necropolis of Torrenueva (3rd-4th century CE, Granada, Spain). Int J Paleopathol 18:26–37
73. Klein S, Avery M, Adams G, Pollard S, Simske S (2014) From scan to print: 3D printing as a means for replication. In: NIP & Digital fabrication conference. Society for imagining science and technology, pp 417–421
74. Porter ST, Roussel M, Soressi M (2016) A simple photogrammetry rig for the reliable creation of 3D artifact models in the field: lithic examples from the Early Upper Paleolithic sequence of Les Cottés (France). Adv Archaeol Pract 4:71–86
75. Massinon M, Dumont B, De Cock N, Salah SOT, Lebeau F (2015) Study of retention variability on an early growth stage herbaceous plant using a 3D virtual spraying model. Crop Prot 78:63–71
76. Secher JJ, Darvann TA, Pinholt EM (2017) Accuracy and reproducibility of the DAVID SLS-2 scanner in three-dimensional facial imaging. J Cranio-Maxillofacial Surg 45:1662–1670
77. Katz D, Suchey JM (1986) Age determination of the male os pubis. Am J Phys Anthropol 69:427–435
78. Savall F, Rérolle C, Hérin F, Dédouit F, Rougé D, Telmon N, Saint-Martin P (2016) Reliability of the Suchey-Brooks method for a French contemporary population. Forensic Sci Int 266:586–5e1
79. Mays S (2014) A test of a recently devised method of estimating skeletal age at death using features of the adult acetabulum. J Forensic Sci 59:184–187
80. Kimmerle EH, Konigsberg LW, Jantz RL, Baraybar JP (2008) Analysis of age-at-death estimation through the use of pubic symphyseal data. J Forensic Sci 53:558–568

81. Spradley MK, Jantz RL, Robinson A, Peccerelli F (2008) Demographic change and forensic identification: problems in metric identification of Hispanic skeletons. J Forensic Sci 53:21–28
82. Brůžek J, Murail P (2006) Methodology and reliability of sex determination from the skeleton. In: Schmitt A, Cunha E, Pinheiro J (eds) Forensic anthropology and medicine: complementary sciences from recovery to cause of death. Humana Press Inc., Totowa, pp 225–242
83. L'Abbé E, Steyn M (2012) The establishment and advancement of forensic anthropology South Africa. In: Dirkmaat DC (ed) A companion to forensic anthropology. Wiley-Blackwell, Malden, pp 626–638
84. Kim J (2017) Understanding population-specific age estimation using documented Asian skeletal samples. In: 86th Annual meeting of the american association of physical anthropologists, p 243

## B. OTHER PUBLICATIONS

**Kotěrová, A.**, Šťovíčková, L., Guyomarc'h , P., Brůžek, J. 2017. Nároky na metody odhadu biologického profilu v období globalizace: současný stav a úkoly pro budoucnost. Slovenská antropologia 20:44–47.

## C. CONFERENCES

**Kotěrová, A.**, Velemínská, J., Cunha, E., Brůžek, J. (2019): Application of the 3D quantitative shape analysis method for age assessment from os coxae in European samples. 88th Annual Meeting of the American Association of Physical Anthropologists, Cleveland, USA.

**Kotěrová, A.**, Velemínská, J., Králík, V., Rmoutilová, R., Friedl, L., Růžička, P., Marchal, F., Brůžek, J. (2019): Evaluation of an impact of different 3D surface scanning protocols on sex and age-at-death assessment from os coxae in bioarchaeology. Society for American Archaeology, Albuquerque, USA.

Zazvonilová, E., **Kotěrová, A.**, Brůžek, J. (2018): Staré ženy pohledem antropologie. Medievistické kolokvium FF UK, Praha, CZ.

**Kotěrová, A.**, Velemínská J., Brůžek J. (2017): Test kvantitativního hodnocení povrchu pubické symfýzy dospělých pro odhad věku dožití podle kostry. Antropologické dny 6th–7th of September, Olomouc, CZ.

Brůžek, J., Dupej, J., **Kotěrová, A.**, Rmoutilová, R., Veleminská, J. (2017): Use of the structured light scanner David SLS-2 for recording auricular surface in 3D and implications for age at death assessment. 86th Annual Meeting of the American Association of Physical Anthropologists, New Orleans, USA.

Guyomarc´h P., **Kotěrová A.**, Šťovíčková L., Brůžek J. (2016): Exigences et réalités des méthodes ostéobiographiques au temps de la mondialisation. XXXII Colloquedu GALF, Toulouse, FR.

Brůžek J., **Kotěrová A.**, Šťovíčková L., Guyomarc´h P. (2016): Nároky na metody odhadu biologického profilu v období globalizace: současný stav a úkoly pro budoucnost. Antropologické dny, Hlohovec, SK.

**Kotěrová, A.**, Velemínská, J., Dupej, J., Brzobohatá, H., Brůžek, J. (2015): Disregarding population specificity: its influence on the sex assessment methods from the tibia. 7th European Academy of Forensic Science conference, Prague, CZ.

# LIST OF FIGURES

# LIST OF TABLES